Essays on Industry Response to Energy and Environmental Policy

Essays on Industry Response to Energy and Environmental Policy

A dissertation presented

by

Richard Leonard Sweeney

to

The Committee on Higher Degrees in Public Policy

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Public Policy

Harvard University

Cambridge, Massachusetts

April, 2015

Dissertation Advisors: *Author:*

Robert Stavins Richard Leonard Sweeney

Ariel Pakes

Essays on Industry Response to Energy and Environmental Policy

Abstract

This dissertation consists of three essays on the relationship between firm incentives and energy and environmental policy outcomes.

Chapters 1 and 2 study the impact of the 1990 Clean Air Act Amendments on the United States oil refining industry. This legislation imposed extensive restrictions on refined petroleum product markets, requiring select end users to purchase new cleaner versions of gasoline and diesel. In Chapter 2, I estimate the static impact of this intervention on refining costs, product prices and consumer welfare. Isolating these effects is complicated by several challenges likely to appear in other regulatory settings, including overlap between regulated and non-regulated markets and deviations from perfect competition. Using a rich database of refinery operations, I estimate a structural model that incorporates each of these dimensions, and then use this cost structure to simulate policy counterfactuals. I find that the policies increased gasoline production costs by 7 cents per gallon and diesel costs by 3 cents per gallon on average, although these costs varied considerably across refineries. As a result of these restrictions, consumers in regulated markets experienced welfare losses on the order of \$3.7 billion per year, but this welfare loss was partially offset by gains of \$1.5 billion dollars per year among consumers in markets not subject to regulation. The results highlight the importance of accounting for imperfect competition and market spillovers when assessing the cost of environmental regulation.

Chapter 2 estimates the sunk costs incurred by United States oil refineries as a result of the low sulfur diesel program. The complex, regionally integrated nature of the industry poses many challenges for estimating these costs. I overcome them by placing the decision to invest in sulfur removal technology within the framework of a two period model and estimate the model using moment inequalities. I find that the regulation induced between \$2.8 and \$3.3 billion worth of investment in order to produce this new fuel. The results highlight the importance

of accounting for sunk costs when evaluating environmental regulation, and suggest that the estimation approach used here might provide a viable way to estimate the sunk costs of other environmental policies.

Chapter 3, coauthored with Hunt Allcott, turns the to retail market for water heaters to study the topic of energy efficiency. We run a natural field experiment at a large nationwide retailer to measure the effects of energy use information disclosure, customer rebates, and sales agent incentives on demand for energy efficient durable goods. We find that while a combination of large rebates plus sales incentives substantially increases market share, information and sales incentives alone each have zero statistical effect and explain at most a small fraction of the low baseline market share. Sales agents strategically comply only partially with the experiment, targeting information at more interested consumers but not discussing energy efficiency with the disinterested majority. These results suggest that at current prices in this context, seller-provided information is not a major barrier to energy efficiency investments. We theoretically and empirically explore the novel policy option of combining customer subsidies with government-provided sales incentives.

Contents

List of Tables

List of Figures

Acknowledgments

I would like to thank my advisors, Joe Aldy, Erich Muehlegger, Ariel Pakes and Rob Stavins, for their constant support and guidance. I am also grateful to my coauthor, Hunt Allcott, from whom I learned an incredible amount, and to the folks at Resources for the Future, especially Dallas Burtraw and Karen Palmer, who introduced me to the subject of environmental economics. Throughout graduate school I have benefited tremendously from my peers at the Kennedy School, HBS and in the economics department, especially participants of the environmental economics and industrial organization lunches. Special thanks to Thomas Covert and Thomas Wollmann, who were always generous with their time and fun to be around, even when research was mainly just "search".

Funding from a Resources for the Future Joseph L. Fisher Dissertation Fellowship, Vicki Norberg-Bohm Fellowship, Joseph Crump Fellowship, and Harvard GSAS Dissertation Completion Fellowship is gratefully acknowledged. The data used in Chapters 1 and 2 were obtained under a researcher information access agreement with the Energy Information Administration, and executed with extensive support from Jacob Bournazian and Joseph Conklin of EIA. The experiment described in Chapter 3 was generously funded by the Sloan Foundation. Our legal agreement does not allow us to identify the Retailer in that chapter, or its employees, but I thank the sales agents and management involved, in particular PC and JW. The views and analysis herein, including any errors, are mine alone.

I am thankful to my brother, Ryan, and my father, Rick, for always being supportive and for keeping me grounded. C, K, M, and J provided much needed perspective and distractions. Most of all I thank my amazing wife, Marina, without whom I truly could not have done this.

This dissertation is dedicated to my mother, Joyce, and her mother, Helen Miranda. All I ever wanted to do was make them proud.

Chapter 1

Environmental Regulation, Imperfect Competition and Market Spillovers: The Impact of the 1990 Clean Air Act Amendments on the US Oil Refining Industry

1.1 Introduction

The US oil refining industry is by far the largest and most sophisticated in the world, producing over $750 billion worth of products in 2012. The 1990 Clean Air Act Amendments imposed a series of new restrictions on refineries, described at the time as the greatest operational challenge the industry had faced since the 1970s.[1] A major component of these regulations was the requirement that gasoline and diesel fuel meet higher environmental standards in certain markets. Areas of the country with severe ozone problems were required to use a new grade of gasoline, called reformulated gasoline (RFG), and highway diesel consumers were required to purchase a new low sulfur grade of distillate, called low sulfur diesel (LSD). These regulated products have since traded at a premium over their unregulated counterparts.

While price increases were expected, it was also widely acknowledged that the cost of producing these new fuels would vary across refineries (NPC 1993). The imposition of fuel content regulations balkanized previously integrated gasoline and distillate markets, with the extent of

[1] Scherr, Richard, G. Allan Smalley Jr., and Michael E. Norman. 5/27/1991. "Refining in the '90s", *Oil & Gas Journal*. Accessed 10/14/2014.

fragmentation varying considerably across regions of the country. This combination of heterogeneous compliance costs and differential access to newly created markets suggests that regulation aimed at reducing pollution externalities could have come at the expense of decreased allocative efficiency. The refining industry is characterized by a relatively small number of firms and enormous barriers to entry, making it a perennial concern for lawmakers and competition authorities (GAO 2004, FTC 2004, 2011). Initial evidence from the introduction of RFG found that the extent of price changes was correlated with the number of suppliers in a market (Brown et al., 2009). However, in-depth study mapping these price changes to refinery costs, profits, and consumer welfare has not been done, partly due to a lack of data availability. Government studies analyzing the impact of regulation on refineries use engineering models and run simulations that assume perfect competition.

In this paper I estimate the variable cost of producing the new fuels mandated under the 1990 Clean Air Act Amendments, as well as the the impact on refined product markets. In order to do this, I obtained access to a rich, previously unused confidential database of refinery-level production decisions. For every refinery in the United States, I observe detailed information on all product outputs, crude inputs, and installed technology, by month, beginning in 1986.

Estimating the impact of these regulations on refineries is complicated by a number of factors that are likely to affect other regulatory settings, particularly those related to energy and environmental policy. The main challenge is a lack of suitable control refineries or markets. Although content regulations varied geographically, supply and demand patterns overlap in way such that every refinery either served a market that became regulated or was linked strategically to one. Bulow, Geanakoplos, and Klemperer (1985) showed that, in oligopolistic settings, a firm's actions in one market can change a competitor's strategies in a second market. It is therefore likely that content regulations not only affected prices and production in directly regulated markets, but also spilled over into unregulated markets as well. The situation is complicated further by the fact that all refineries are multiproduct firms. Faced with increased costs or competition in one product market, such as gasoline, refineries could have responded by increasing relative production of other products, such as jet fuel. Accounting for this cross-product adjustment margin is important for understanding the net costs of the policy.

I overcome these challenges, as well as the more fundamental challenge of costs not being observed, by estimating a structural model of refinery decisions which directly incorporates each of these dimensions. Employing methods first introduced by Rosse (1970), the empirical approach is to estimate the costs of producing reformulated gasoline and low sulfur diesel by comparing refiners' willingness to supply gasoline and distillate across seemingly similarly profitable situations before and after the regulation. I assume constant elasticity of demand and Cournot competition, and develop a multiproduct marginal cost function for each refinery. After recovering the cost structure for the industry, I then simulate counterfactuals with the fuel content restrictions removed to calculate the price and welfare effects of the policy.

The main finding of the paper is that reformulated gasoline increased refinery costs by 7 cents per gallon and low sulfur diesel increased costs by 3 cents per gallon on average, although these costs varied considerably across refineries. I find that the demand for petroleum products at the wholesale level is more elastic than would be surmised from end-user consumption patterns alone. As a result, while refineries producing reformulated gasoline saw markups increase in those markets, average refinery profits from 1995 to 2003 were 8 percent lower than they would have been without content regulation. Contributing to this decline in profits was a reduction in margins in conventional product markets, as refineries which found it difficult to produce the new fuels reallocated output elsewhere. Thus, while consumers in regulated markets experienced welfare losses on the order of \$34 billion during this period, this loss was partially offset by gains of \$14 billion dollars among consumers in markets not subject to regulation.

Several papers have looked at the impact of gasoline content regulation on prices. Muehlegger (2006) estimates the relationship between regional content regulations and gasoline price spikes, finding that content regulation contributed to price spikes in California, Illinois and Wisconsin. Brown et al. (2008) use detailed weekly wholesale price data from 1994 to 1998, and compare prices across matched regulated and unregulated cities in the years immediately before and after gasoline content regulations went into effect. They find that regulation increased prices by 3 cents per gallon on average, but that this effect varied considerably across cities depending on their degree of market isolation. Chakravorty, Nauges, and Thomas (2008) use data from 1995 to 2002 to estimate the impact of "boutique" gasoline standards on state-level wholesale prices. They model differences in gasoline standards across states as being endogenous to the concentration of refineries in each state, and find that OLS estimates understate the effect of regulation on prices. Although no papers have estimated the cost of low sulfur diesel, Zhang (2011) estimates the cost of the subsequent switch from low sulfur to ultra-low sulfur diesel in 2006.[2] This paper extends this literature by incorporating refinery-level data and separately estimating costs and price effects. I show that costs vary significantly across refineries, and directly relate this cost heterogeneity to changes in market power and markups.

This paper also contributes to a growing empirical literature at the intersection of industrial organization and environmental economics (Millimet, Roy, and Sengupta 2009). Ryan (2012) estimates a dynamic structural model of the Portland cement industry to assess the impact of the 1990 Clean Air Act Amendments on that industry. He finds that focusing on static prices and profits alone generates negative cost estimates, but that the sign is reversed once changes in fixed costs are incorporated. In a subsequent paper, Fowlie, Reguant, and Ryan (2014) use this model to assess the relationship between alternative carbon pricing policies and product market distortions in this setting. They find that policies which fail to account for strategic firm behavior generate social welfare losses. In this paper, I show that allowing for imperfect competition and accounting for strategic interdependence between markets is important understanding the full effects of environmental regulation. Previous reduced form studies of the Clean Air Act have implicitly calculated the gross effect of regulation on regulated versus unregulated areas or firms, leaving the net impact ambiguous. In this paper, I account for intra-country shifting by modeling refinery decisions explicitly and then simulating policy counterfactuals to recover the net national effect of fuel content regulations. The results highlight the importance of allowing for market spillovers, and demonstrate how detailed firm-level data can be combined with assumptions about producer behavior to recover regulatory impacts in settings where the program evaluation framework is not applicable.

The remainder of the paper proceeds as follows. Section 1.2 provides a brief overview of the refining industry and describes the relevant environmental regulation. Section 1.3 introduces and summarizes the data. Section 1.4 develops a structural model of refinery behavior and

[2]In addition to these papers, several authors have empirically studied the oil refining industry. Berman and Bui (2001) find that productivity at refineries in Los Angeles increased during a period when they were subject to stringent point source emission regulations. Considine (2001) describes a structural model of markup pricing under joint production at refineries. Hendricks et al (2007) introduce a model of bilateral oligopoly to study mergers in this industry. Hastings and Gilbert (2007) find evidence of raising rivals costs after Tosco's acquisition of Unocal. Chesnes (2014) studies the impact of refinery outages on product prices and refinery investment. Further discussion of how this paper relates to prior structural models of the industry is delayed until Section 3.4, after more background on the industry is provided.

describes the estimation procedure. Section 1.5 presents the main results of the paper, and Section 1.6 concludes.[3]

1.2 Institutional background

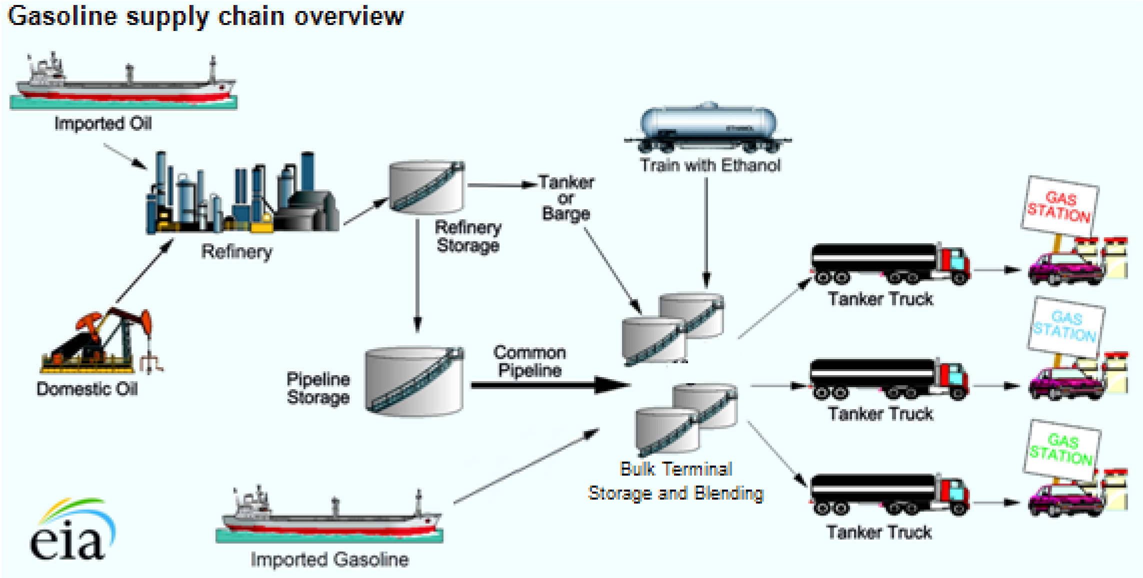
To provide intuition for the structural model specified below, in this section I give a brief overview of the refining industry and relevant environmental regulation. A key feature of the industry is that all refineries produce multiple products and production is an inherently joint process. The efficiency of this process varies across refineries, driven by differences in crude oil input quality and installed technology. Geography is also important, as pipelines are by far the cheapest way to transport products, and every refinery is not connected to every state by pipeline. This paper focuses on fuel content regulation, which mandated that certain drivers purchase reformulated gasoline or low sulfur diesel. The share of local gasoline and distillate markets covered under the regulation varied across states, and, due to the incompleteness of the pipeline network, this meant that some refineries were more affected by the regulations than others.

1.2.1 The oil refining industry

Refineries lie at the middle of the US transportation fuel supply chain (Figure 1.1). Crude oil is extracted upstream, domestically or abroad, processed at a refinery, and then shipped out via pipeline or barge to wholesale terminals, where it is distributed by truck for local consumption. Crude oil as it comes out of the ground is a mixture of different length hydrocarbons, ranging from short hydrocarbons, which roughly correspond to butane and gasoline, to long hydrocarbons, which correspond to asphalt and tar. At the most basic level, oil refining consists of separating crude into streams of differing densities using heat and a complex series of catalytic processes. The "lighter" end products, which include gasoline, diesel and propane, are typically of much higher value. So, all else equal, a refiner tries to maximize the amount of light outputs produced from a given amount of crude oil.

[3]Each section begins with a brief summary of the key points for the time constrained reader.

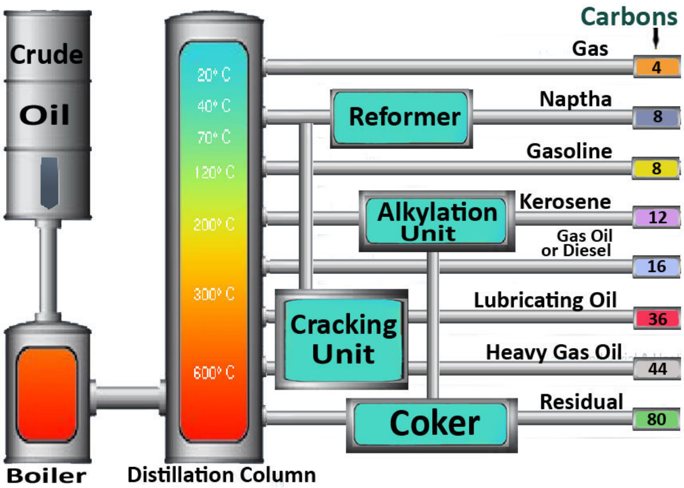
Figure 1.1: Petroleum Product Supply Chain



The finished product mix obtainable from a barrel of crude oil is a function of the type of crude that is used and the type of processing capital the refinery has installed. Crude oil properties vary across oil fields, and are typically described by two characteristics, API gravity and sulfur content. API gravity, denominated in degrees, is a measure of crude density, which dictates the relative proportions of light and heavy oils that can be separated out during simple distillation. Heavier crudes contain a relatively larger share of long hydrocarbons, which translates into a larger share of heavy end products. Sulfur content, denominated in parts per million (ppm), was historically of interest because it causes corrosion in metals and other processing complications. More recently, environmental regulations have set caps on the amount of sulfur end products can contain. Thus, light "sweet" (i.e. low sulfur) crudes, like West Texas Intermediate (WTI), are typically the most valuable.

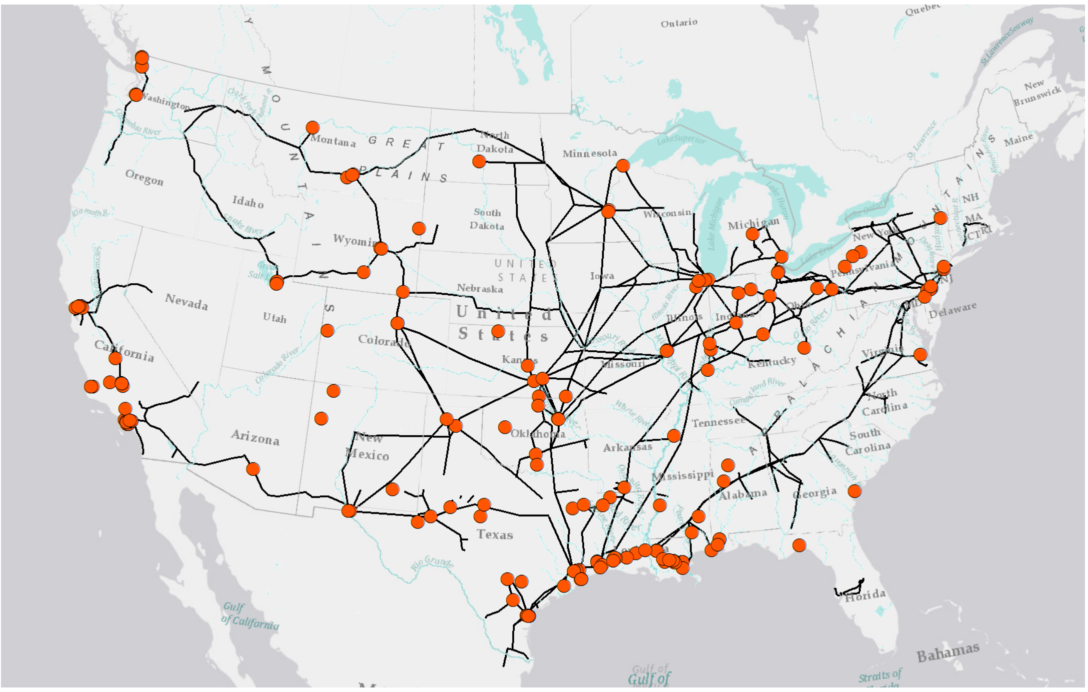
While API gravity and sulfur content map fairly directly into the proportions of refined outputs obtained from simple distillation, modern refineries in the United States are much more complex operations (Figure 1.2). It is therefore more useful to think of gravity and sulfur content as determining the amount of processing necessary to transform a given type of crude into a particular end product mix. The most basic refining technology is the distillation tower, which separates crude into different density streams by slowly boiling it. These streams can then be sent through any number of secondary processes, collectively referred to as "upgrading" capacity, which increase the yield of higher value light end-products. Finally, there are a host of technologies that can be employed at the end of the process to remove pollutants and impurities. From an environmental perspective, the most important of these processes is hydrotreating, which removes sulfur. Refineries in the US range from simple "topping" operations, which do not have any upgrading capacity and have light yields of less than 50 percent, to the most complex refineries in the world, where the amount of upgrading capacity exceeds distillation capacity and light yields routinely top 90 percent. However, even at the most sophisticated operations, it is impossible to transform an entire barrel of crude into gasoline or diesel, making refining an inherently joint production process.

Figure 1.2: Modern Refinery Configuration



The final important dimension of differentiation in the industry is geography. Figure 1.3 presents a map of refineries and refined product pipelines in the continental United States. Historically, refineries were set up where crude oil was easily available. As a result, US refining is relatively geographically concentrated and not particularly well correlated with the location of end-users. This map actually understates the extent of regional concentration, as the refineries in the Gulf Coast are much larger than other areas, giving this region almost 50 percent of total US capacity. In order to balance the location of supply with demand, an extensive pipeline distribution system evolved over the course of the 20th century to transport refined products to local markets. These pipeline routes are typically unidirectional, with product generally flowing north from refining centers to populations centers. There are two key points to note about this pattern. East Coast refining capacity is only about a third of its consumption. As result, it receives around 50 percent of its supply from the Gulf Coast and accounts for almost all of the refined product imports in the US. On the opposite extreme, the West Coast is shielded from European imports as well as domestic imports via pipeline from the rest of the country. It therefore relies on refineries in the region for almost all of its consumption.

Figure 1.3: Refinery and Pipeline Locations



1.2.2 Environmental regulation

The refining industry is one of the largest sources of air pollution in the United States, contributing both directly through emissions generated during the refining process and indirectly through the combustion of petroleum products at end sources. As a result, refineries have been subjected to considerable environmental regulation over the last half century, stemming primarily from the Clean Air Act of 1963 and subsequent amendments (Table 1.1). Direct emissions were covered under regulations which placed increased oversight on polluting facilities located in counties which did not meet newly established National Ambient Air Quality Standards (NAAQS) (Greenstone 2002). The 1970 Clean Air Act Amendments also targeted indirect refinery emissions by permitting regulation of the chemical composition of refined petroleum products. This authority was used to phase out lead in gasoline starting in 1975, and to reduce volatile organic compounds (VOCs) and other ozone precursors by imposing Reid vapor pressure (RVP) limits in summer months beginning in 1989.[4] The 1990 Clean Air Act Amendments (CAAA) marked the most significant regulation of indirect refinery emissions to date, requiring oxygenate be added to gasoline in some markets and a new reformulated version of gasoline be supplied in others, depending on the region and season.[5] The CAAA also imposed strict sulfur limits on highway diesel fuel.

Table 1.1: **Summary of Environmental Regulation**

Regulation	**Description**	**Dates**
RVP Limits	Limits on gasoline in summer months (May 1 - September 15)	Phase 1: 1989 - 1991 Phase 2: 1992 - present
Oxygenated Gasoline	Required oxygenate be blended into gasoline in severe CO non-attainment counties (November - February)	1992 - present
Federal RFG	Content and performance limits on gasoline in severe ozone non-attainment counties	Phase 1: 1995 - 2000 Phase 2: 2000 - present
CARB Gas	RFG with additional restrictions for CA	Phase 1: 1992 Phase 2: 3/96 Phase 3: 4/03
Low Sulfur Diesel	Highway diesel capped at 500 (ppm)	10/93 - 6/06
Tier 2 Low Sulfur Gas	Average gasoline sulfur content set to 30 ppm	Phased in: 2004 - 2006
Ultra Low Sulfur Diesel	Highway diesel capped at 15 (ppm)	Phase-in: 6/06 - 5/10 Binding 6/10

In this paper, I focus on reformulated gasoline (RFG) and low sulfur diesel (LSD) from 1994 to 2003. Of the regulations introduced by the 1990 Clean Air Act Amendments, these were the two that called for direct alterations to the refining process.[6] In 1999, new sulfur limits were announced for gasoline and highway diesel, which phased in starting in 2004 and 2006

[4]Reid vapor pressure is a measure of the propensity of gasoline to evaporate. RVP regulation was implemented in two phases, affecting summer months starting in 1989 and 1992. See Auffhammer and Kellogg (2011) for more information on RVP regulation.

[5]See Muehlegger (2004) for a thorough review of gasoline content regulation.

In addition to these refinery specific programs, many refineries were also covered under New Source Review. This program subjected large polluting facilities located in non-attainment areas to undergo additional regulatory review before making investments or plant alterations. Approximately 70% of refining capacity was located in regions that were out of attainment. Although this designation restricted investment opportunities, it did directly affect operations conditional on installed capital. In future work I plan to study the long run impacts of this program on refinery investment by nesting the static profit structure estimated in this paper within a dynamic refinery investment game.

[6]The other CAAA content regulation, oxygenated gasoline, simply involved blending in oxygenate at the end of the refining process. This was often done by a third party downstream.

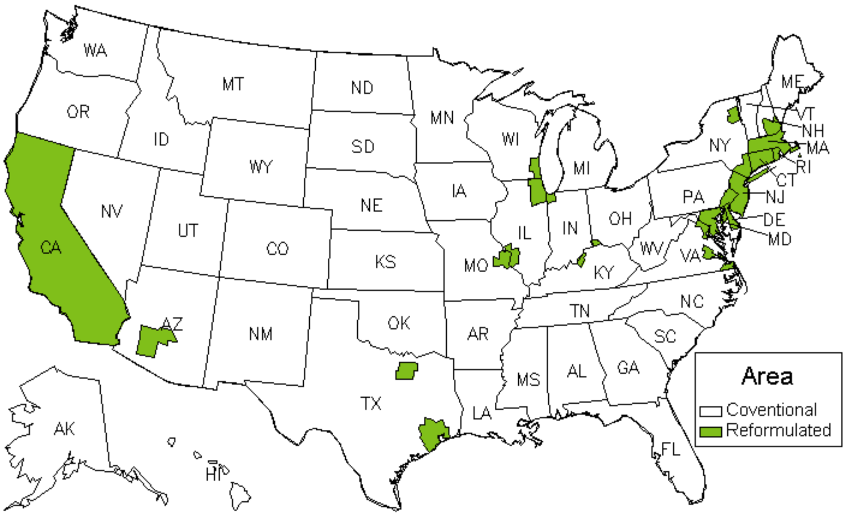
respectively. These new programs allowed for flexibility in compliance. Specifically, firms could average sulfur reductions across refineries, buy or sell excess reductions from other firms, and bank early compliance for future credit. Unfortunately, actual sulfur content levels and modes of compliance for each refinery are not publicly available.[7] Incorporating this phase-in period and flexibility requires a substantial extension to the model specified below, and is left for future work.

1.2.2.1 Reformulated Gasoline

The 1990 Clean Air Act Amendments mandated the adoption of reformulated gasoline in nine large metropolitan markets with severe ozone pollution. RFG is gasoline manufactured to reduce the amount of smog forming particles and toxic pollutants released into the air during combustion. The program was implemented in two phases, coming online in 1995 and 2000. Both phases set minimum oxygenate levels of 2 percent and capped benzene levels at 1 percent of volume. Phase I also required a 15 percent reduction in toxic air pollutants relative to conventional gasoline, and this reduction was increased to 25 percent in Phase II. In addition to these year round requirements, RFG imposed stricter VOC limits during summer ozone season (June 1 - September 15), again mandating reductions of 15 and 25 percent in Phases I and II. Finally, Phase II added a year round NOx requirement of 5.5 percent. While only nine areas were required to use RFG, other areas with moderate ozone pollution were allowed to opt into the program. Today, RFG is used in 17 states and the District of Columbia, making up 30 percent of US gasoline consumption (Figure 1.4). In March 1996, California and Arizona adopted a more stringent version of RFG, called CARB gasoline, which imposed tighter seasonal VOC limits and an 80 percent reduction in sulfur content.

[7] A Freedom of Information Act request to obtain realized trades and compliance methods was denied by the EPA.

Figure 1.4: Counties Requiring Reformulated Gasoline

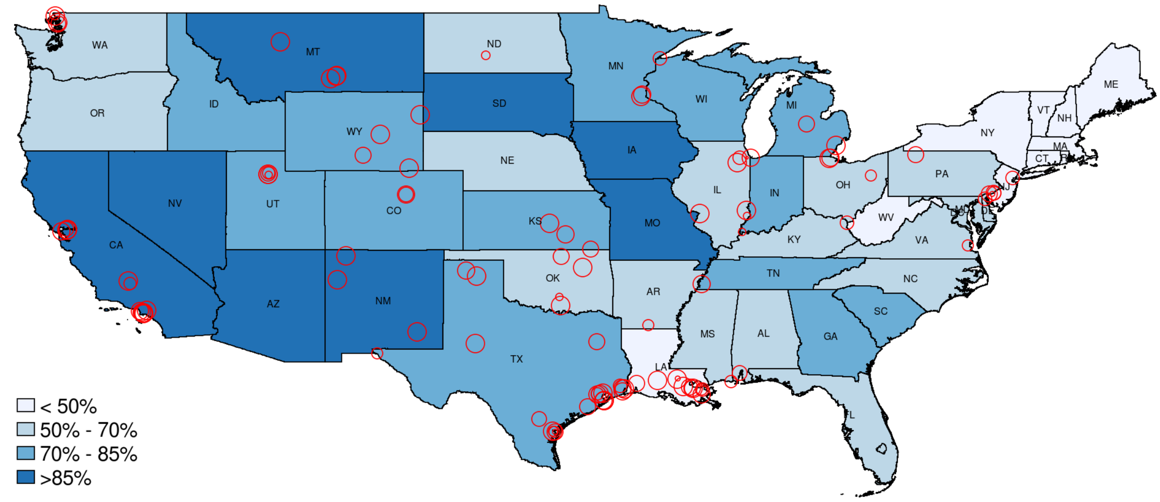


In order to qualify as reformulated, gasoline had to meet specific composition and emission performance criteria. Rather than an exact formula, there were many different ways to satisfy the criteria, allowing refineries the flexibility to make different tradeoffs based on the economics of their operations. As a result, producing RFG is not associated with any specific piece of refinery equipment or characteristic, although, at a minimum, refineries had to reconfigure their operations at the beginning of the program. For many refineries, the extent of reconfiguration required was too costly, and they opted not to participate in this market.

1.2.2.2 Low Sulfur Diesel

In order to facilitate new particulate emissions standards on heavy duty diesel engines, the CAAA capped the sulfur content of on-highway diesel at 500 ppm starting in October 1993. Distillate fuel oil is a general classification of relatively heavier petroleum products used for domestic heating, industrial burners, and compression in ignition engines. Diesel fuel is distillate fuel burned in diesel engines. "On-highway" diesel fuel is diesel fuel used by trucks and passenger cars, whereas "off-highway" diesel is used in farms, construction, and marine vessels. Distillate fuel oil is thus primarily categorized by end use, rather than physical properties. This distinction is important, because the new distillate regulations imposed by the CAAA only applied to one of these categories. Home heating oil and other similar distillates were not required to meet the standard, which affected 46 percent of distillate demand, and around 8 percent of total petroleum demand at the time of enactment. Although the delineation was not as stark as RFG, there is still considerable heterogeneity in the fraction of distillate that was regulated under the program across states, driven by differences in relative demand by end-use. Figure 1.5 presents a map with the each state's share of distillate consumption that was low sulfur after the regulation began. In the northeast, substantial amounts of distillate are used for home heating oil, meaning that less than 50 percent of those markets were covered under the new regulation.[8] At the other extreme, in the southwest, over 90 percent of distillate sales after 1994 were highway diesel.

Figure 1.5: Share of Post-94 Distillate Sales that are Low-Sulfur



NOTE: Circles proportional to the average fraction of distillate that is low-sulfur for each refinery

Producing this new low sulfur distillate was a significant achievment for refineries, as the national average sulfur content at the time was 3000 ppm (Lidderdale 1993). In contrast to RFG, LSD production was largely determined by the installed capacity of hydrotreating or "desulfurization" capacity. According to Lidderdale (1993), refineries with catalytic hydrocracking units may be able to reconfigure them to remove some sulfur, but, otherwise, LSD production would be largely limited by desulfurization capacity. Smaller, less sophisticated refineries were therefore particularly vulnerable to this regulation, and, in an effort to compensate for this, were given SO2 credits for sulfur removed from diesel. California again adopted a slightly more stringent version of the regulation, imposing a 10 percent aromatics limit on highway diesel in addition to the 500 ppm sulfur cap.

[8]In other parts of the country, off-highway distillate is mainly used in farming equipment or marine vessels.

1.3 Data description and summary

In this section I describe and summarize the data obtained from EIA. Detailed input and production data is observed for every refinery in the United States. Sales are observed at the firm-product-state level, and the demand side of these transactions is primarily comprised of intermediaries which reallocate products to end consumers. The sample is restricted to 124 large, non-specialty refineries, 20 of which exit between 1986 and 2003. There is considerable variation in the level of participation in reformulated gasoline and low sulfur diesel markets across these refineries, driven largely by geographic location. Despite this exogenous variation in regulatory exposure, reduced form regressions of refinery RFG and LSD production on gasoline and distillate productivity fail to identify any significant costs associated with these regulations.

1.3.1 Description of confidential EIA data

Through a confidential data request I obtained several data sets on refinery operations from the Energy Information Administration (EIA), each described in appendix Table A.1. The main data come from survey EIA-810, which contains monthly information on all inputs and outputs for every refinery in the United States from 1986 to 2012. Importantly, gasoline is reported as being conventional or reformulated, and distillate is broken out by sulfur level. The average API gravity and sulfur content of crude inputs is also reported, along with the amount of distillation capacity in operation at the start of each month. This monthly distillation capacity information is supplemented with annual data from survey EIA-820, which records the capacity of every refinery unit, including all upgrading and desulfurization capacity, at the start of each year.[9]

This refinery-level data is combined with several firm-level data sets. Survey EIA-782A is a census of monthly state-level sales by every firm which owns a refinery in the United States. Refiners report sales in the state where the transfer of title occurred, regardless of where that product is ultimately consumed. Both the volume sold and the price are reported, broken out by sales to end users (retail) and sales for resale (wholesale). Survey EIA-782C is a census of all "prime suppliers", which includes firms that own refineries as well as large importers and marketers. 782C asks respondents to only report sales for which they are the final supplier into

[9]EIA-820 was not collected in 1996 and 1998. For these years, I interpolate capacity for each refinery based on the reported values from the adjacent years

a state and the fuel is going to be consumed within state. The 782C data does not contain price and does not break volume sold into retail and wholesale.

The distinction between the 782A and 782C data is important for understanding the demand system specified in this paper. Firms record 782A sales in the state where the transfer of product occurred and 782C sales in the state where the product is ultimately consumed. In this sense, the quantity reported on 782A reflects the total quantity demanded *from refineries* for transfer in each state, while 782C reflects the quantity demanded *by end-users* in a given state from all distributors.[10] In this paper, I assume the relevant demand curve facing refineries is reflected in the 782A data. There are several reasons for this. First, survey 782A records the location and price of transactions where they occurred, and it is the price at and shipping costs to this location, rather than where the purchaser ultimately transports the product, which presumably matters to refineries. Second, the 782A data breaks sales down by retail and wholesale channels, and I use this information in estimation below. Finally, as the purpose of the 782C data is to measure state-level consumption, many transactions are excluded from the data in order to avoid double counting.[11]

One major limitation of the data is that distribution is observed at the firm level, as opposed to the refinery level. I attempt to overcome this by assuming that firms minimize transportation costs when serving end markets. I obtained GIS maps of the US refined product pipeline system and waterways suitable for petroleum transportation from EIA, along with GIS coordinates of each refinery. Following Muehlegger (2006), costs for transporting petroleum products by pipelines, barges and trucks of 2, 4.5 and 30 cents per gallon per thousand miles are taken from estimates presented before the Federal Trade Commission (Jacobs 2002). For each refinery, I calculate the least cost method of serving each state. I then minimize each firm's total cost of meeting its observed state-level sales from the 782A data, subject to the observed 810 output at each of its refineries.

The final confidential data set comes from survey EIA-14, which contains average crude oil prices, including cost of delivery, at the firm-PADD-month level starting in 2004. For earlier

[10] In general, states with many refineries, located at key points in the pipeline network tend to have quantities in 782A which exceed prime supplier volumes in 782C, whereas states that are net importers of refined products the opposite is true. EIA (2009) discusses the differences between the surveys in great detail.

[11] For example, if a refinery sells product to a distributor who then resells it in another state, quantity from this first transaction would not appear in the 782C data.

years I use predicted crude delivery price as a function of benchmark crude spot prices, region dummies, regional domestic crude first purchase prices, and crude prices binned by API gravity (Appendix B). This confidential data is supplemented with annual state-level population data from the Census Bureau, petroleum product taxes and vehicle registrations from the Federal Highway Administration, and monthly weather data from the National Oceanic and Atmospheric Administration.

1.3.2 Summary Statistics

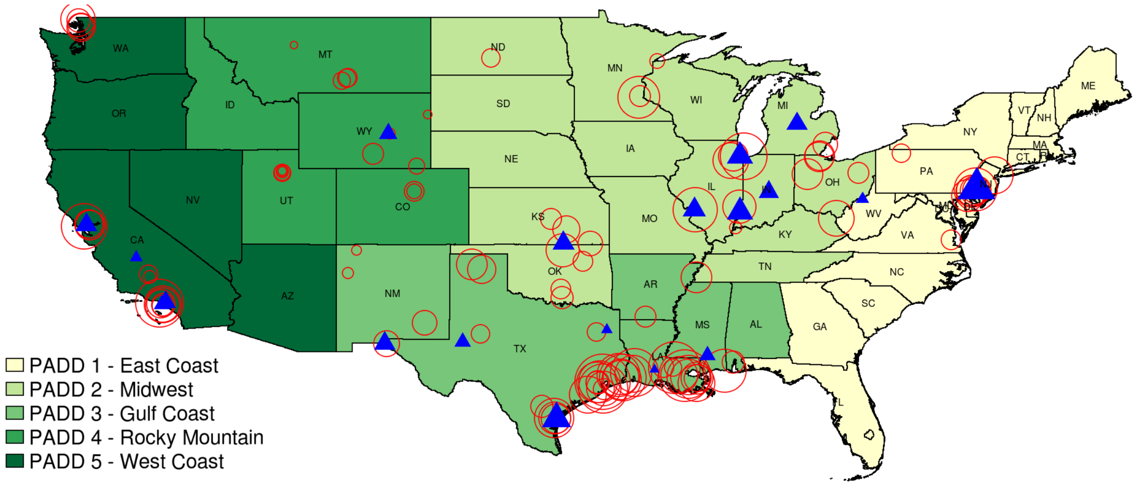
Table 1.2 presents refinery summary statistics from the data. The sample is restricted to a pre period (1986-1992) and a post period (1995-2003). 1993 and 1994 are omitted because LSD and RFG, respectively, appear in production data for months before they are tracked in the consumption data, so price and destination are not observed. Refineries are grouped by Petroleum Administration Defense Districts (PADDs), which is a commonly used geographic aggregation dating back to World War II. Figure 1.6 provides a map of the regions, as well as a proportional representation of the refineries in the sample. 220 refineries appear at some point in the data. Small refineries, with less than 10,000 barrels per day operating capacity, and specialty refineries, with less than 50 percent light production, are excluded from the analysis.

Table 1.2: Summary Stats By Region

	East Coast (PADD 1)		Midwest (PADD 2)		Gulf Coast (PADD 3)		Rockies (PADD 4)		West Coast (PADD 5)	
# Refineries										
1986	11		30		47		16		20	
1995	9		29		44		15		18	
2003	9		24		40		15		16	
Average Yields (%)										
Gasoline	42.4	(7.9)	51.1	(6.)	45.0	(7.8)	44.5	(8.1)	43.0	(12)
Distillate	24.5	(4.7)	23.6	(6.9)	22.8	(6.8)	26.2	(4.9)	17.9	(7.1)
Jet Fuel	5.2	(4.6)	4.6	(4.5)	7.3	(5.5)	4.7	(4.2)	9.4	(7.2)
Capacity (KBbl/cd)	133	(69)	113	(83)	151	(119)	35	(16)	121	(76)
Upgrading Capacity (%)	50.8	(25)	49.7	(12)	58.0	(40)	42.6	(15)	67.2	(35)
API Gravity (degrees)	33.7	(5.1)	35.3	(4.6)	34.4	(6.)	35.5	(5.6)	26.2	(5.)
Crude Sulfur (%)	0.8	(0.7)	0.9	(0.6)	1.1	(0.8)	0.9	(0.8)	1.1	(0.3)
# Producing RFG	8		9		24		0		13	
# Producing LSD	9		25		39		15		16	
Wholesale Gasoline Prices										
Pre-1993 Conventional	1.25	(.15)	1.18	(.16)	1.13	(.15)	1.19	(.17)	1.26	(.16)
Post-1994 Conventional	1.01	(.2)	1.02	(.22)	0.96	(.2)	1.11	(.2)	1.15	(.21)
Post-1994 Reformulated	1.13	(.21)	1.17	(.22)	1.03	(.2)	-	-	1.23	(.23)
Wholesale Distillate Prices										
Pre-1993 High-Sulfur	1.06	(.18)	1.06	(.17)	1.01	(.17)	1.12	(.19)	1.05	(.2)
Post-1994 High-Sulfur	0.90	(.21)	0.93	(.2)	0.85	(.2)	1.03	(.21)	0.95	(.22)
Post-1994 Low-Sulfur	0.92	(.21)	0.95	(.21)	0.89	(.2)	1.03	(.21)	1.02	(.21)

Notes: Sample restricted to refineries with more than 10KBbl/cd and at least 50% light yield. All prices in real (2013) dollars per gallon.

Figure 1.6: Refinery Size and Exits by PADD



NOTE: Blue triangles indicate refineries which exited by 2003. Points are proportional to ending distillation capacity.

Of the 124 refineries included in the sample, 9 exited between 1986 and 1995, and another 11 exited by 2003.[12] Most of these exits represent the tail end of an industry restructuring after deregulation. In 1981, the US removed domestic oil price controls, causing 101 relatively small or inefficient refineries to exit the industry between 1981 and 1985. This trend of exits continued, albeit at a much slower pace, through the 1990's. Figure 1.7 reports the number of operating refineries and total industry distillation capacity in each year. Although a large number of refineries exited the industry, they were relatively small operations, and the amount of operating capacity lost was more than offset by expansions at refineries that stayed. No new refineries have been built in the US since the 1970's. Despite this, market structure varies substantially over the course of this period, driven by a wave of mergers and acquisitions in the 1990s (Table 1.3).

[12]In four of these exits, the refinery was sold to a nearby refinery and integrated into that refinery's operations. Of the 11 refineries that exit post 1994, two in California and two Illinois produced RFG. The two in CA exited in the summer of 1995, citing inability to comply with the stricter CARB phase II specifications (FTC 2006). Premcor shut down its Illinois refineries in 2001 and 2002 citing high capital costs to meet upcoming sulfur regulations.

Figure 1.7: Distillation Capacity Changes

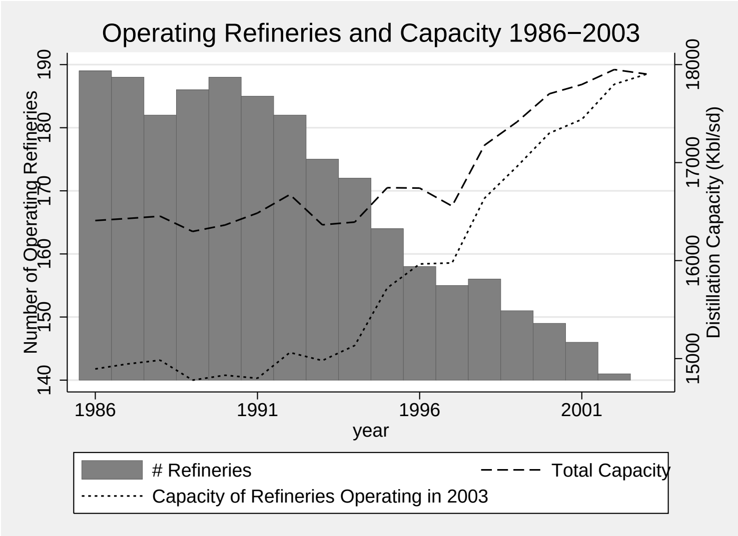


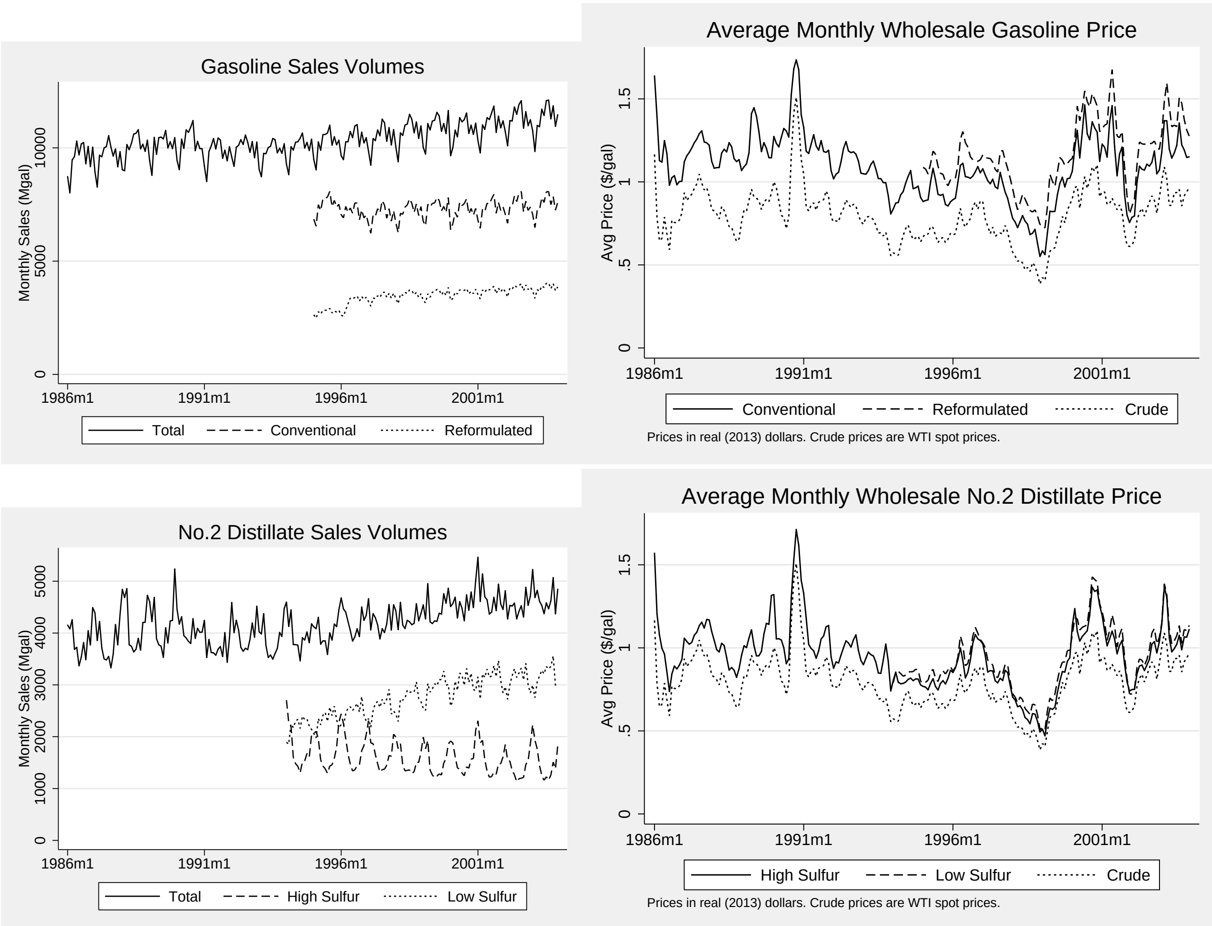
Table 1.3: Operable Refineries and Sales 1986 - 2003

Year	Operable Refineries	Refinery Sales
1986	216	15
1987	219	9
1988	213	21
1989	204	5
1990	205	4
1991	202	9
1992	199	2
1993	187	2
1994	179	4
1995	175	6
1996	169	7
1997	164	19
1998	160	40
1999	159	11
2000	158	14
2001	155	29
2002	153	6
2003	149	8

Notes: A detailed genealogy of US refiners can be found at `http://www.eia.gov/finance/genealogy/`.

Figure 1.8 displays national price and quantity trends for gasoline and distillate during the sample. Total gasoline volumes were relatively constant, while total distillate sales have been trending upward since the late 1980's due to increasing demand for highway diesel. Prices of all products varied substantially over this time period, driven mainly by global crude price movements (Choinard and Perloff 2008).

Figure 1.8: Price and Quantity Trends



Turning back to the refinery summary statistics in Table 1.2, the average yields of gasoline, distillate and jet fuel are 45 percent, 23 percent, and 6 percent respectively, with much more variation across refineries within each region than across regions. The Gulf Coast has the largest refineries, with an average ability to process 150,000 barrels of crude oil per day. The West Coast has the most sophisticated refineries, with an average ratio of upgrading capacity to distillation capacity of over 67 percent. The West Coast also uses the heaviest crude oil during this period, while refineries in the Rocky Mountain region are both the smallest and the least sophisticated.

The next two rows in the table report the number of refineries in the sample that produce RFG and LSD by region. Only 54 of the 111 refineries operating after 1994 produce reformulated gasoline, where as 104 refineries produce low sulfur diesel. On the intensive margin, there is heterogeneity across refineries in the share of each regulated product produced (Figures 1.9 and 1.10). As was discussed in Section 1.2.2, geography is a major determinant of market access in this industry, and regions differed in how affected they were by content regulation. Table 2.1 formally tests the extent to which this influenced refinery production of these newly created products. The dependent variable in these regressions is the share of a refinery's gasoline that is RFG, the share of a refinery's distillate production that is LSD, and the change in desulfurization capacity between 1990 and 1996. The independent variable is the average share of RFG and LSD post-1994 in states that each refinery was serving prior to 1990, when the regulations were announced. In all three regressions the share of a refinery's pre-1990 markets which were subsequently subject to CAAA content regulation significantly determines its post-1994 production shares of the new fuels.

Figure 1.9: Average RFG Share by Refinery

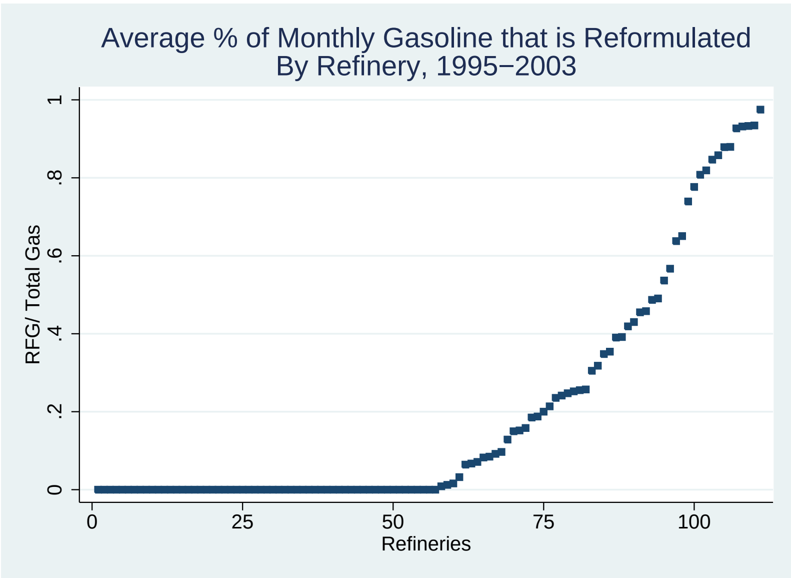


Figure 1.10: Average LSD Share by Refinery

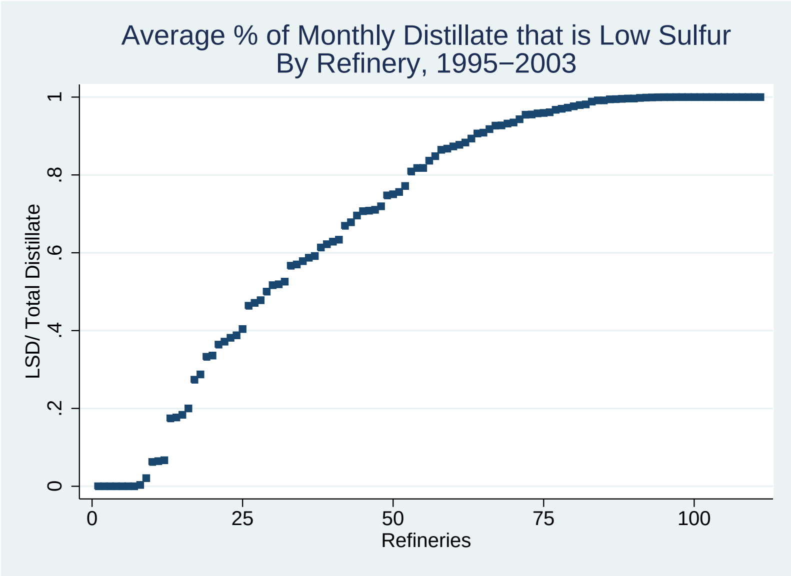


Table 1.4: Geographic Determinants of RFG and LSD

	Post 94 % RFG	Post 94 % LSD	Δ Desulf Cap 90-96
% States Served Pre 1990	0.934***	1.341***	0.900***
	(0.0673)	(0.421)	(0.235)
N	122	122	121
F	192.5	10.14	14.70

Notes: The independent variable in each regression is the share of each refinery's pre-1990 markets which subsequently became regulated post-1994. The dependent variable in the first two models is the share of each refinery's gasoline and distillate production which is regulated in the post period. The dependent variable in third column is the change in desulfurization capacity per unit of distillation capacity.

The bottom section of Table 1.2 reports average wholesale prices for each product and region. Conventional gasoline prices were higher on the coasts prior to the introduction of RFG. In the post 1994 period, RFG prices were around 10 cents per gallon higher than conventional prices. Low sulfur diesel was 6 cents per gallon more than high sulfur distillate in PADD 5, and only 1 cent per gallon more in PADD 4, which had the highest distillate prices to begin with. Appendix Table A.2 reports average prices by state, along with the average Herfindahl–Hirschman Index (HHI) for each market. The FTC generally considers markets with an HHI in excess of 0.15 to be moderately concentrated. At the state level, 11 gasoline and distillate markets met this criterion in the pre period, and 16 conventional markets, 10 RFG markets, and 19 LSD and

HSD markets were above this threshold in the post period.

1.3.3 Reduced form results

Before turning to the structural model, it is useful to look at the reduced form relationship between RFG and LSD production and refinery productivity. There are two channels through which the regulation could have affected operations: by altering the amount of gasoline or diesel extracted from each barrel of crude, or by changing the amount of crude the refiner chose to process. One measure of refinery productivity is therefore the ratio of gasoline or distillate output to crude distillation capacity, which is the product of these two channels. Table 1.5 presents results from the following regression,

$$q_{ijt}/k_{it} = \beta X_{it} + \alpha_j \theta_{ij} + \nu_p + \mu_i + \gamma_t + \epsilon_{it} \tag{1.1}$$

Where q_{ijt} is the number of gallons of product j (i.e. gasoline or distillate) produced at refinery i in month t, and k is installed monthly distillation capacity, denominated in gallons of crude oil processable per month. θ_{ij} is the fraction of gasoline or distillate that is reformulated or low-sulfur respectively. ν_p is an indicator for the post-1994 period interacted with PADD dummies, and μ and γ are refinery and time period fixed effects. X includes the installed upgrading capacity and crude API gravity at each refinery each period.

Table 1.5: Impact of Content Regulation on Output

	Average RFG %	RFG %	RFG % - IV
Gasoline Output / Capacity	0.0852***	0.0527*	-0.0139
	(0.0323)	(0.0271)	(0.0528)
N	18595	18595	18560

	Average LSD %	LSD %	LSD % - IV
Distillate Output/ Capacity	0.0528***	0.0340***	0.0445***
	(0.0126)	(0.00964)	(0.0109)
N	18595	18595	18560

Notes: All regressions include API gravity, upgrading capacity for major downstream units, indicators for the post-1994 period interacted with PADD dummies, and refinery and time dummies. The first stage F-stat for the IV models is 13.48 for RFG and 74.11 for LSD. Standard errors are clustered at the refinery level.

Three separate specifications for θ_{ij} are run for both gasoline and distillate. In the first column, θ_{ij} is the average RFG and LSD share for each refinery from 1995 to 2003 (Figures 1.9 and 1.10). A 10 percent increase in RFG or LSD shares is associated with 0.8 and 0.5 percent higher yields of gasoline and diesel per unit of capacity. The second column uses monthly variation in θ_{ij}. Within-refinery deviations in RFG and LSD shares are again associated with higher gasoline and distillate productivity. Of course, θ_{ij} is chosen by the refinery, and could be correlated with other refinery unobservables or with residual demand shocks to relative RFG and LSD demand. The third column presents IV results where I instrument for θ_{ij} using the pre-1990 market share variable from Table 1.4 as well as gasoline and highway diesel taxes. The distillate regression also uses heating degrees days as an instrument, which should increase relative demand for high-sulfur distillate. After instrumenting, the RFG coefficient is slightly negative, but not significant, while the LSD coefficient is still large and significant. There does not appear to be strong evidence that producing RFG or LSD reduces refinery productivity, and, in the case of LSD, appears to increase the amount of distillate obtained from a given capital stock and crude type.[13]

1.4 A model of refinery operations

In this section I develop a structural model of the refining industry that extends the existing literature by simultaneously incorporating joint production, capacity constraints, and imperfect competition. In order to run policy counterfactuals, I need to recover all of the parameters governing refinery behavior, not just those related to content regulations. I specify a multiproduct cost function and use a logit transformation to incorporate the joint nature of the refining process. Costs are not directly observed, and are instead inferred from market clearing decisions under assumptions about producer behavior, following Rosse (1970). Firms compete simultaneously in quantities in each state, and estimation is based off first order conditions which equate marginal revenue in each end market to a centralized marginal cost of production for each refinery. Identification comes from an extensive set of seasonal, temporal and geographic demand shifters, as well as changes in refinery ownership and capacity shares, which vary the

[13]A log-log specification, rather than ratios, yields similar results, as does including PADD-year or PADD-time dummies.

infra-marginal quantity each refinery internalizes when making production decisions.

1.4.1 Existing literature

Three authors have previously estimated structural models of the US refining industry using aggregated data on refinery operations. Muehlegger (2006) estimates a marginal cost function for refined products, but assumes separable production with perfect substitutability. With refinery-level production data, I am able to observe behavior that is much more consistent with joint production. Despite the fact the relative prices of end products vary by plus or minus 40 percent during this period, there is a not a single observation where a refinery produces only gasoline or diesel. Chesnes (2009) estimates a dynamic model of refinery investment, but assumes that product yields are fixed across refineries and over time. As was shown in Table 1.2 and Figure 1.8, product yields vary significantly across refineries and time periods. Since one potential response for a refinery facing content regulation was to alter its production mix, it is important in this paper to explicitly incorporate this margin.

Zhang (2011) estimates a multiproduct production function, but assumes perfect competition and treats refinery operating levels as exogenous and identical across refineries in a PADD. With refinery-level data I am able to observe that the average coefficient of variation of refinery utilization rates within a PADD-month is 0.15. Moreover, variation in utilization rates is correlated with market power. A regression of monthly refinery utilization on the controlling firm's share of total capacity in the PADD along with refinery and time fixed effects returns a coefficient of -0.327 (0.149), indicating that a 10 percent increase in regional market share is associated with a 3.2 percent reduction in marginal willingness to operate (see Appendix C for further details). Again, incorporating market power is particularly important in this paper, as it is possible that markups for the new products differed as well as costs. In what follows, I specify a model that includes a multiproduct cost function, incorporates endogenous refinery-level heterogeneity in yields and utilization, and allows for costs to differ from marginal prices based on market shares.

1.4.2 Structural model

Firms face a constant elasticity of demand curve for each product j and end market m:

$$\ln Q_{jm}(\alpha) = \alpha_{0jm} + \alpha_j \ln p_{jm} + \epsilon^D_{jm} \tag{1.2}$$

Competition is assumed to take place at the wholesale level. In the 782A data, 87 percent of gasoline sales and 83 percent of distillate sales are sales for resale. At this level, products are essentially homogeneous. Although branded gasoline often contains additives which garner a price premium at the pump, these are added at the wholesale terminal by the purchasing party. When shipped, refined products of a particular type commingle, with purchasers often unaware of which refinery the product was produced at.

Products are shipped to markets from refineries i at transportation costs τ, resulting in total post-shipping revenues:

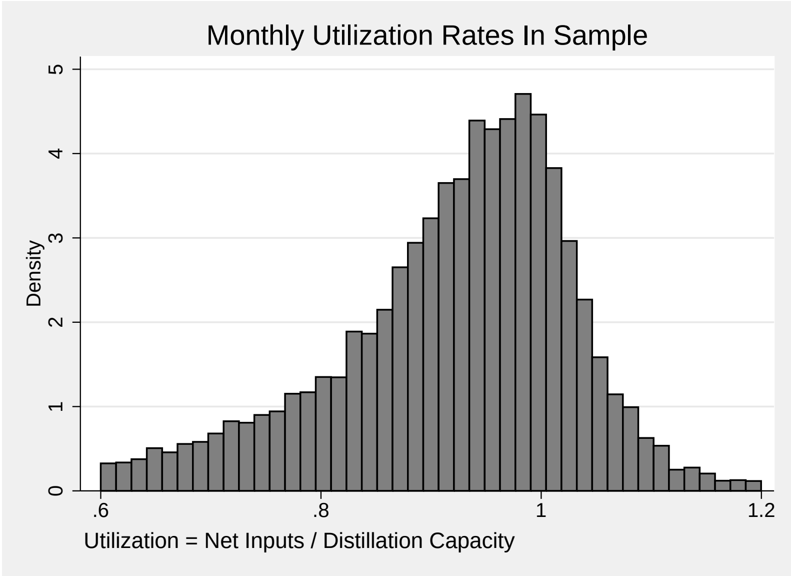
$$Rev_i(\mathbf{q_i}) = \sum_j \sum_m (p_{jm}(Q_m) - \tau_{im}) q_{ijm} \tag{1.3}$$

These revenues are generated at the expense of a single centralized production cost for each refinery:

$$Cost_i(\mathbf{q_i}, c_i; \beta) \quad = \quad \sum_j (\beta_j + \beta_{jj} \frac{q_{ij}}{c}) q_{ij} + (\beta_c + p_c) c_i + f(\frac{c_i}{k_i}) \tag{1.4}$$

The cost function has two components, product specific processing costs and general operating costs. β_j reflects the constant marginal cost of producing each product, while β_{jj} represents the increasing difficulty of extracting higher yields for each product. In addition to these product specific costs, for each gallon of crude c processed, the refinery pays constant marginal operating costs β_c, crude oil costs p_c, and convex utilization costs $f(\frac{c}{k})$, where k is the amount of installed distillation capacity. Refineries report capacity as the number of gallons of crude processable under "normal" operating conditions, rather than the maximum processable amount. As a result, monthly utilization rates routinely exceed 100 percent, although the distribution drops off sharply after that point (Figure 1.11).

Figure 1.11: Observed Utilization Rates



Content regulations RFG (r) and LSD (l) enter the cost function by shifting the product specific marginal processing costs of gas and distillate, mc_g and mc_d,

$$mc_{ir} = \beta_g + \beta_{gg}\frac{q_{ig}}{c_i} + \beta_r = mc_{ig} + \beta_r \tag{1.5}$$
$$mc_{il} = \beta_d + \beta_{dd}\frac{q_{id}}{c_i} + \beta_l = mc_{id} + \beta_l \tag{1.6}$$

Under this specification, β_r and β_l represent the dollar per gallon cost increase of RFG and LSD relative to conventional gasoline and distillate.

As was discussed in Section 1.2.1, refining is an inherently joint production process. Operations are centered around "production runs", where a refinery sets the amount of each end product to extract from each gallon of crude and then decides how much crude to process. Define the yield of gasoline and distillate from a given gallon of crude as $Y^g = \frac{q_g}{c}$ and $Y^d = \frac{q_d}{c}$, and the share of each that satisfies content restrictions as $\theta_r = \frac{q_r}{q_g}$ and $\theta_l = \frac{q_l}{q_d}$. The cost function becomes,

$$Cost_i(\mathbf{Y_i}, \theta_\mathbf{i}, c_i; \beta) = c_i \left[(\beta_g + \beta_{gg} Y_i^g + \beta_r \theta_{ir}) Y_i^g + (\beta_d + \beta_{dd} Y_i^d + \beta_l \theta_{il}) Y_i^d + \beta_o Y_i^o \right] + c_i(\beta_c + p_c) + f(\frac{c_i}{k_i}) \quad (1.7)$$

Where $Y^o = (1 - Y^g - Y^d)$ is the refinery's outside option from increasing gasoline and distillate yields, which includes the all other refined petroleum products, such as jet fuel, residual fuel oil, asphalt and propane.

Firms are assumed to compete simultaneously in quantities. Combining equations (1.7) and (1.3), and summing over all refineries I_f owned by firm f yields the firm's optimization program:

$$\begin{aligned} Max_{\{q_{ijm}, Y_i, \theta_i, c_i\}} \pi &= \sum_{i \in I_f} Rev_i(\mathbf{q}_i) - Cost_i(\mathbf{Y_i}, \theta_\mathbf{i}, c_i; \beta) \quad (1.8) \\ s.t. \ \sum_m q_{ijm} &\leq Y_i^j c_i \\ \sum_j Y_i^j &= 1 \end{aligned}$$

Despite relaxing many assumptions made in the previous literature, the model is still limited in several respects. Vertical integration in either direction is not captured. Upstream, this implies that firms are price takers in the crude oil market. Approximately 55 percent of refining capacity in the sample is owned by independent refiners with no upstream operations. Downstream, the model assumes that each refinery's incentive to supply a market is fully captured in the wholesale price. In reality, some wholesale sales are going to parties the refiner has a contractual obligation with, or that compete against its downstream arm.[14] However, information on these relationships is not available for all products or for the entire sample. Finally, demand is assumed to be static, ignoring the demand side's ability to smooth purchases across time through storage and refineries' consideration of inventories when setting quantities each period (Borenstein and Shepard 2002).

[14] Gilbert and Hastings (2005) find evidence that vertical integration had a significant impact on wholesale gasoline prices following Tosco's acquisition of Unocal's West Coast assets in 1997.

1.4.3 Estimation

Estimating equations are based on the first order conditions of (1.8). Two assumptions are made to simplify the problem, which contains hundreds of first order conditions. First, for all products, the Karush-Kuhn-Tucker conditions imply that marginal revenue, net of shipping costs, must be equal in expectation for all markets served by each refinery.

$$mr_{ijm} = p_{jm}(1 + \frac{s_{fjm}}{\alpha_j}) - \tau_{im} = \frac{\partial Cost_i(\mathbf{Y_i}, \theta_\mathbf{i}, c_i; \beta)}{\partial q_{ij}} \quad \forall q_{ijm} > 0 \tag{1.9}$$

Where s_{fjm} is firm f's market share and α_j is the demand elasticity parameter from (1.2). Let mr_{ij} be the expected marginal revenue of each product across all markets served by refinery i, assuming optimal allocation.

Second, rather than working in fractions with constraints, I use the logit transformation to convert yield choices into a continuous unbounded state space:

$$Y^g = \frac{e^{\delta_g}}{1 + e^{\delta_g} + e^{\delta_d}} \quad ; \quad Y^d = \frac{e^{\delta_d}}{1 + e^{\delta_g} + e^{\delta_d}} \tag{1.10}$$

Conceptually, δ_j can be thought of as encompassing all of the effort and resources a refinery allocates towards producing product j. This specification also incorporates the inherent multiproduct nature of the process, by imposing the logit assumption that the effect of increased effort towards product j is proportional to the yields of products j and k,

$$\frac{\partial Y^j}{\partial \delta_j} = Y^j(1 - Y^j) \quad ; \quad \frac{\partial Y^k}{\partial \delta_j} = -Y^k Y^j$$

Incorporating these assumptions reduces the refiner's problem to five choices variables for each month: how much refining effort to direct towards gasoline and diesel (δ_g, δ_d), the share of each to convert into RFG and LSD (θ_r, θ_l), and how many gallons of crude to process (c). Each decision variable is assumed to have an associated private cost shock ϵ^S known to the refinery at the time of production but unobserved to the econometrician. Under this formulation, the cost function becomes,

$$Cost_i(\delta_{\mathbf{i}}, \theta_{\mathbf{i}}, \mathbf{c_i}; \beta) \;=\; c_i\left[(\beta_g + \beta_{gg}Y_i^g + \beta_r\theta_{ir})Y_i^g + (\beta_d + \beta_{dd}Y_i^d + \beta_l\theta_{il})Y_i^d + \beta_o Y_i^o\right] \quad (1.11)$$
$$+c_i(\beta_c + P_c) + f(\frac{c_i}{k_i}) + c_i\epsilon_{ic}^S + k_i(\delta_{ig}\epsilon_{ig}^S + \delta_{id}\epsilon_{id}^S + \theta_{ir}\epsilon_{ir}^S + \theta_{il}\epsilon_{il}^S)$$

Each of the product specific cost shocks is assumed to enter additively and scale with capacity across refineries. The interpretation here is that these shocks pertain to configuring the refinery to generate a specific output mix each production run, but do not directly affect marginal crude input decisions conditional on yields. Fixed costs are not identified in this model, and assumed to be zero.[15] Differentiating with respect to each choice variable yields five first order equations to be estimated simultaneously:

$$\text{(FOC1)} \quad \frac{d\pi}{d\delta_{ig}} = \frac{c_i}{k_i}\left[[mr_{ig}(1-\theta_{ir}) + mr_{ir}\theta_{ir} - mr_{io}]Y_{ig}^g + [mr_{id}(1-\theta_{il}) + mr_{il}\theta_{il} - mr_{io}]Y_{ig}^d\right]$$
$$-\frac{c_i}{k_i}\left[(\beta_g + \beta_{gg}Y_i^g + \beta_r\theta_{ir})Y_{ig}^g + (\beta_d + \beta_{dd}Y_i^d + \beta_l\theta_{il})Y_{ig}^d\right] - \epsilon_{ig}^S = 0$$
$$\text{(FOC2)} \quad \frac{d\pi}{d\delta_{id}} = \frac{c_i}{k_i}\left[[mr_{ig}(1-\theta_{ir}) + mr_{ir}\theta_{ir} - mr_{io}]Y_{id}^g + [mr_{id}(1-\theta_{il}) + mr_{il}\theta_{il} - mr_o]Y_{id}^d\right]$$
$$-\frac{c_i}{k_i}\left[(\beta_g + \beta_{gg}Y_i^g + \beta_r\theta_{ir})Y_{id}^g + (\beta_d + \beta_{dd}Y_i^d + \beta_l\theta_{il})Y_{id}^d\right] - \epsilon_{id}^S = 0$$
$$\text{(FOC3)} \quad \frac{d\pi}{dc_i} = [mr_{ig}(1-\theta_{ir}) + mr_{ir}\theta_{ir} - mr_{io}]Y_i^g + [mr_{id}(1-\theta_{il}) + mr_{il}\theta_{il} - mr_{io}]Y_i^d$$
$$+mr_{io} - (\beta_g + \beta_{gg}Y_i^g + \beta_r\theta_{ir})Y_i^g - (\beta_d + \beta_{dd}Y_i^d + \beta_l\theta_{il})Y_i^d$$
$$-\beta_c - p_c - f'(\frac{c_i}{k_i}) - \epsilon_{ic}^S = 0$$
$$\text{(FOC4)} \quad \frac{d\pi}{d\theta_{ir}} = \frac{c_i}{k_i}Y_i^g\left[(mr_{ir} - mr_{ig}) - \beta_r\right] - \epsilon_{ir}^S = 0$$
$$\text{(FOC5)} \quad \frac{d\pi}{d\theta_{il}} = \frac{c_i}{k_i}Y_i^d\left[(mr_{il} - mr_{id}) - \beta_l\right] - \epsilon_{il}^S = 0$$

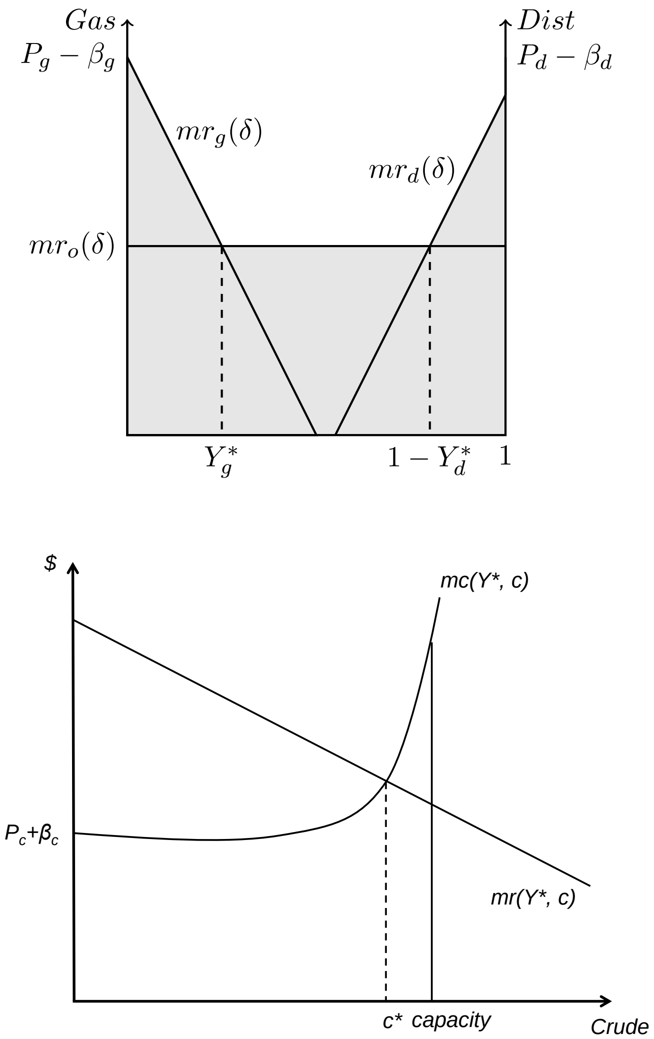
Where $Y_k^j = \frac{\partial Y^j}{\partial \delta_k}$, and $f'(\frac{c}{k})$ is the marginal utilization cost per gallon of crude processed.

Figure 1.12 provides a graphical representation of the firm's problem. The top panel represents the yield choice facing the refiner. When choosing δ_g and δ_d, the marginal revenue from increasing yields at a given crude level is equated to the marginal revenue of the outside

[15] Unlike electric power generators, refineries always operate except for scheduled maintenance (every 3-5 years) or an unscheduled disruption, such as a fire, which prohibits operation.

option. Content regulations shift the intercept of the marginal net revenue curves for gasoline and distillate by $(P_r - P_g - \beta_r)$ and $(P_l - P_d - \beta_l)$ respectively. The slope of the curves will also change to the extent that market shares in regulated markets differ from unregulated markets. The shaded gray area reflects the total marginal revenue gained at optimal yields Y^* for a given level of crude inputs c. The second panel represents the refiner's decision of how much crude to process, where $mr(Y^*, c)$ indexes this optimal marginal revenue for every possible level of operation. Refineries face increasing utilization costs as they approach capacity, while the price of crude oil and other constant marginal costs shift the point at which these utilization costs intersect marginal operating revenue.

Figure 1.12: Refiner's Problem



All of the marginal cost intercepts are modeled as linear functions of refinery characteristics, $\beta_j = \sum_i \beta_{jk} X_{jki}$. For both gas and diesel, X_j includes a constant, API gravity and API gravity squared, crude oil sulfur content, the ratio of upgrading capacity to distillation capacity, and the interaction between API gravity and upgrading capacity. X_g also includes indicators for summer months (May - September) post-1989 and 1992 to capture summer gasoline RVP restrictions. β_o includes the share of the outside option that is jet fuel, which is the most valuable product not explicitly modeled here. All other products in the outside option are assumed to trade at the price of residual fuel oil, which is a benchmark bottom of the barrel petroleum product tracked in the EIA data. Operating costs β_c are assumed to be zero, beyond refinery and time fixed effects, and marginal utilization costs are modeled as a cubic B-spline with knots placed at quartiles of the utilization distribution.

LSD costs are modeled as follows,

$$\begin{aligned} \beta_{li} &= \beta_{l0} + \beta_{l1} PctUpgrading + \beta_{l2} PctDesulf + \beta_{l3} PctDesulf^2 + \beta_{l4} Sulfur \\ &\quad + \beta_{l5} PctDesulf * Sulfur + \beta_{l6} API + \beta_{l7} PctDesulf * API + \beta_{l8} CA + \beta_{l9} SmallRefinery \end{aligned} \tag{1.12}$$

Where $PctUpgrading$ and $PctDesulf$ are the ratios of installed upgrading and desulfurization capacity to total distillation capacity, $Sulfur$ is the sulfur content of crude oil, CA is an indicator for the state of California, which imposed stricter diesel limits, and $SmallRefinery$ is an indicator for refineries eligible to receive SO2 credits for producing LSD.

As was discussed above, unlike LSD, RFG capability was not a function of any observable refinery technology. I therefore estimate refinery-specific RFG costs,

$$\begin{aligned} \beta_{ri} &= \beta_{r0i} + \beta_{r1} RFG1Summer + \beta_{r2} RFG2 + \beta_{r3} RFG2Summer \\ &\quad + \beta_{r4} CARB + \beta_{r5} CARBSummer + \beta_{r6} MTBE \end{aligned} \tag{1.13}$$

Where β_{ri} is a refinery dummy for all refineries with positive RFG production, $RFG1$ and $RFG2$ indicate the phases of the RFG program, and $Summer$ indicates summer months, which involved added restrictions. $CARB$ indicates reformulated gasoline sold in California or Arizona after March 1996, and $MTBE$ is an indicator for whether the refinery was able to use MTBE as an oxygenate in making RFG. For the early years of the program, MTBE was the preferred mode of satisfying the oxygenate requirements of RFG. Beginning in the late 1990s, there was increasing public concern that MTBE was in fact carcinogenic. In response to this, a number of states banned MTBE between 2000 and 2006, at which point a federal ban was enacted (Anderson and Elzinga 2012).

I estimate four demand functions from equation (1.2): gasoline demand, pre-1993 distillate, post-1994 high sulfur distillate and low sulfur diesel.[16] Each demand equation contains state-month dummies to account for variations in seasonality and time dummies. These four equations are estimated jointly with the supply side first order conditions, resulting nine equations and errors $\epsilon = (\epsilon^D, \epsilon^S)$.

Estimation proceeds via 2-stage GMM by jointly minimizing $E(\epsilon' Z)$, where $Z = \{Z_D, Z_S\}$ is a set of instruments that are assumed to be uncorrelated with demand and supply errors respectively (Hansen 1982). Z_D includes regional crude prices, regional refinery capacity concentration, and pipeline outages, which should be correlated with prices but unrelated to demand shocks. Z_S includes end-market population, weather, and fuel taxes, all of which shift demand but should not alter refinery production costs. Z_S also includes firm level capacity share, the number of refineries operating in each region, and regional capacity concentration, which vary considerably during the sample and shift the residual demand curve facing each refinery. Finally, Z_S also includes month dummies. Both gasoline and distillate exhibit considerable seasonality. Figure 1.13 shows that while gasoline demand increases in the summer, distillate demand is relatively higher in the winter. Figure 1.14 breaks distillate sales down further to reveal that highway diesel demand is actually slightly higher in the summer, while demand for high sulfur distillate, which is largely used for heating oil, is over twice as high in the winter compared to the summer. This seasonal variation in relative demand helps pin down the convexity associated increasing gasoline and distillate yields, as well as the costs of converting distillate into low-sulfur diesel.

[16] Although distillate sales are broken out into diesel and fuel oil prior to 1993, Marion and Muehlegger (2008) show demand for these two products was jointly determined, as diesel consumers sought to evade taxes by purchasing untaxed distillate intended for off-highway use. Concurrent with the introduction of low sulfur diesel, the government mandated that non-highway distillate be marked with a dye to prevent illegal sales.

Figure 1.13: Seasonality in Gasoline and Distillate Demand

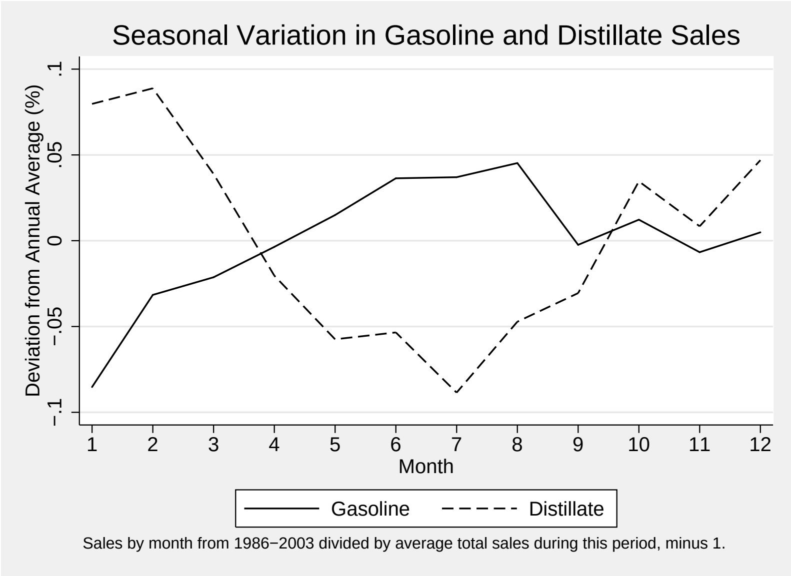
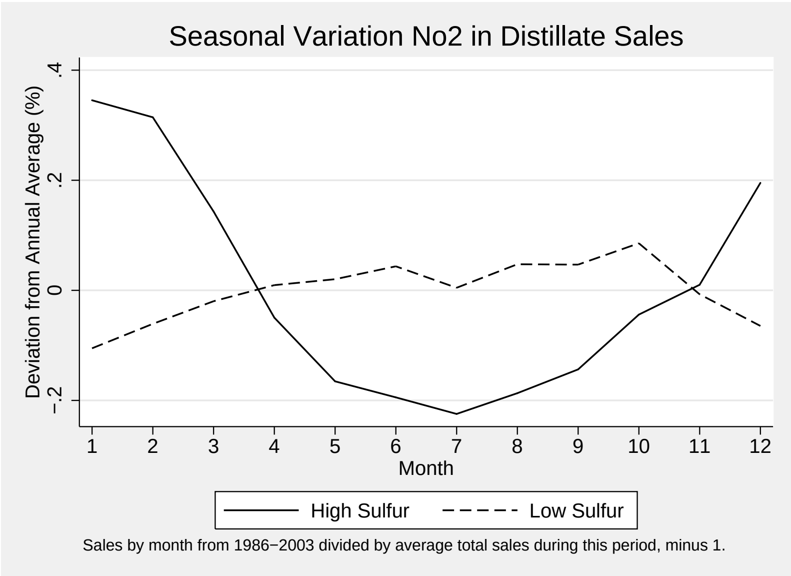


Figure 1.14: Seasonality in Distillate Demand by Sulfur Content



All parameters enter the supply equations linearly, and $\epsilon_g^S, \epsilon_d^S$, and ϵ_c^S each contain refinery and year dummies. These equations therefore use two sources of variation to identify the costs of content regulation. First, with refinery and year fixed effects, (FOC1-FOC3) compare refineries' willingness to supply gas and distillate, and to operate, before and after content regulation. As was discussed in section 1.2.2, refineries were differentially exposed to content regulation based on their proximity to regulated markets. In addition to this cross-period comparison, equations

(FOC4-FOC5) compare refineries' willingness to tradeoff between regulated and unregulated products within a given month.

1.5 Results

This section presents the main results from the paper. Cost function estimation returns intuitive coefficients on crude quality and refinery technology, and significant heterogeneity in the marginal cost of producing RFG and LSD across refineries. Wholesale demand estimates are more elastic than commonly studied end-use petroleum product demand elasticities, and the precision of these estimates increases substantially when the supply side and demand side are estimated simultaneously. In order to recover the industry equilibrium price and quantity effects of content regulation, I simulate counterfactual market outcomes for the entire United States with fuel content restrictions removed. Under this counterfactual, prices in RFG and LSD markets are 6 and 3 percent lower on average, and prices in conventional gasoline and distillate markets are 2 and 1 percent higher. Although markups in RFG markets are higher with the restrictions, refineries experience decreased profits on average due to decreased markups and volumes on other products.

1.5.1 Cost function estimates

Cost function estimates are presented in Table 1.6. Model 1 uses wholesale prices in mr_{ij} for all sales, and Model 2 uses a weighted average of wholesale prices and sales through company owned retail outlets for each firm. As expected, using a higher quality crude (one with a high API gravity and low sulfur content) is associated with lower costs of gasoline and distillate production. Costs are also reduced by having more upgrading capacity, but the interaction of these two is positive, reflecting the fact that higher quality crude needs less processing. The slope coefficient on gasoline is slightly higher than for distillate, with both positive and significant, reflecting the fact that it is costly to increase product yields conditional on capital and crude quality. Jet fuel is estimated to cost 37 cents per gallon more than the outside option. Utilization costs are estimated to be essentially flat over the most of the distribution, increasing sharply once a refinery's crude inputs exceeds 100 percent of installed capacity. In Model 2, retail sales of gasoline and distillate have an effect equivalent to increasing marginal costs by 16 and 4 cents per gallon respectively, although this includes both retailing costs and the average marginal profit from retailing.

Table 1.6: Cost Function Estimates

	Model 1		Model 2	
Gasoline	**Est**	**SE**	**Est**	**SE**
Constant	0.379	(0.051)	0.582	(0.059)
API Gravity	0.176	(0.234)	-1.026	(0.273)
API Gravity^2	-1.243	(0.310)	0.233	(0.352)
% Upgrading	-0.181	(0.039)	-0.311	(0.044)
% Upgrading * API	0.480	(0.120)	0.913	(0.132)
Crude Sulfur	-0.003	(0.002)	-0.003	(0.002)
% Retail			0.161	(0.004)
RVP 1989	0.084	(0.004)	0.087	(0.004)
RVP 1992	-0.005	(0.003)	-0.002	(0.003)
Yield	0.147	(0.017)	0.137	(0.017)
Distillate				
Constant	0.241	(0.053)	0.415	(0.054)
API Gravity	0.706	(0.233)	-0.303	(0.240)
API Gravity^2	-1.536	(0.305)	-0.137	(0.314)
% Upgrading	-0.048	(0.037)	-0.132	(0.038)
% Upgrading * API	0.064	(0.112)	0.303	(0.114)
Crude Sulfur	0.007	(0.002)	0.001	(0.002)
% Retail			0.043	(0.002)
Yield	0.092	(0.021)	0.087	(0.022)
Other				
Jet fuel	0.376	(0.009)	0.373	(0.009)
Utilization = .85	-0.209	(0.040)	-0.083	(0.036)
Utilization = .94	-0.166	(0.036)	-0.060	(0.032)
Utilization = .99	-0.165	(0.037)	-0.055	(0.033)
Utilization = 1.2	0.106	(0.049)	0.164	(0.047)
Regulation Costs				
Average RFG Cost	0.096	(0.008)	0.071	(0.009)
Average LSD Cost	0.029	(0.004)	0.033	(0.004)
Elasticities				
Gasoline	1.61	(0.040)	1.32	(0.032)
Pre-94 Dist	2.00	(0.082)	2.05	(0.092)
HSD	3.18	(0.100)	4.34	(0.208)
LSD	2.09	(0.074)	2.22	(0.089)
Psuedo R2				
Gasoline Moment	0.66		0.67	
Distillate Moment	0.62		0.61	
Crude Moment	0.63		0.67	
RFG Moment	-0.03		0.15	
LSD Moment	-0.06		-0.06	
Gasoline Demand	0.87		0.88	
HSD Demand Pre-94	0.92		0.92	
HSD Demand Post-94	0.83		0.82	
LSD Demand	0.94		0.94	

Notes: All models contain 20,227 observations. The first three moments (FOC1-FOC3) contain refinery dummies and year dummies. All demand equations contain state-month dummies and time dummies. Robust standard errors presented.

Table 1.7 presents the estimated costs of RFG and LSD in detail, along with projected costs from the EPA and the National Petroleum Council.[17] The average intercept β_{r0i} across the 54 refineries which produced RFG during the sample is 12.3 and 9.3 cents per gallon in the models excluding and including retail prices. Phase I summer restrictions are not estimated to significantly increase RFG costs beyond nationwide summer RVP limits. Costs are actually estimated to be slightly lower during Phase II, although the summer component of Phase II, which was the most stringent addition, is large and significant. Summer restrictions in California increase costs by a similar amount. The inability to use MTBE is found to substantially increase RFG costs by 3.6 to 3.8 cents per gallon. In combination, the results indicate that RFG was 7.1 to 9.6 cents per gallon more expensive than conventional gasoline on average. When weighted by quantity of RFG produced, the implied increase in refining costs is 6.2 to 8.0 cents per gallon. All but the unweighted Model 1 estimates fall within the range predicted by EPA, and below that predicted by the NPC. Figure 1.15 plots the average RFG cost for each refinery with positive RFG production in the sample. Each point is a refinery, and they are sorted by cost. The cost estimates vary considerably across refineries, ranging from -2.7 cents per gallon all the way up to 23 cents per gallon, 16 cents higher than the mean.

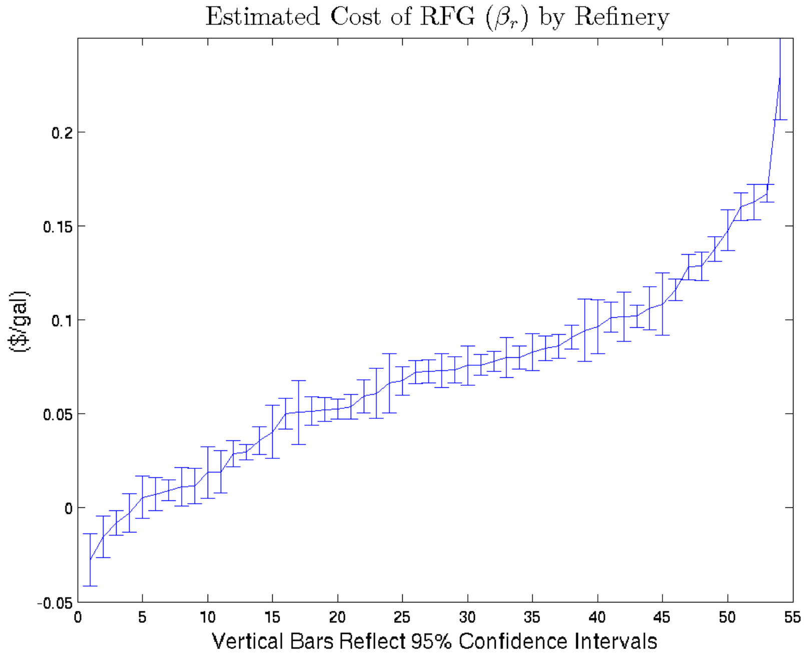
[17] The NPC is a federal advisory committee to the Secretary of Energy made up of petroleum industry executives. The purpose of the NPC is to advise and make recommendations to the Secretary of Energy (FTC 2006).

Table 1.7: Content Regulation Costs

	Model 1		Model 2	
Reformulated Gasoline	**Est**	**SE**	**Est**	**SE**
Constant	12.3	(0.48)	9.3	(0.56)
RFG 1 Summer	-0.7	(0.33)	0.0	(0.36)
RFG 2	-1.7	(0.12)	-0.7	(0.15)
RFG 2 Summer	3.5	(0.24)	3.1	(0.29)
CARB	-0.6	(0.59)	0.7	(0.68)
CARB Summer	4.6	(0.36)	3.4	(0.41)
MTBE	-3.6	(0.25)	-3.8	(0.29)
Average RFG Cost	9.6	(0.80)	7.1	(0.93)
Wgt. Avg. RFG cost	8.0		6.2	
EPA Estimates				
Phase 1	4.8-7.8			
Phase 2	8.6			
NPC Estimates				
Phase 1	8.6			
Phase 2	11.5			
Low Sulfur Diesel				
Constant	2.3	(1.27)	0.7	(1.43)
Upgrading Capacity	-0.4	(0.30)	-2.8	(0.35)
% Desulfurization	3.8	(2.03)	7.2	(2.29)
% Desulfurization ^ 2	1.3	(0.60)	1.1	(0.63)
Crude Sulfur	0.7	(0.24)	1.1	(0.27)
% Desulfurization * Sulfur	-1.9	(0.33)	-2.1	(0.38)
API Gravity	11.9	(2.93)	17.2	(3.38)
% Desulfurization * API	-28.9	(4.05)	-32.8	(4.67)
CA	2.4	(0.26)	2.4	(0.28)
Small Refinery	1.1	(0.15)	1.5	(0.17)
Average LSD Cost	2.9	(0.36)	3.3	(0.42)
Wgt. Avg. LSD cost	2.2		2.6	
EPA Estimate	4.3			
NPC Estimate	6.8			

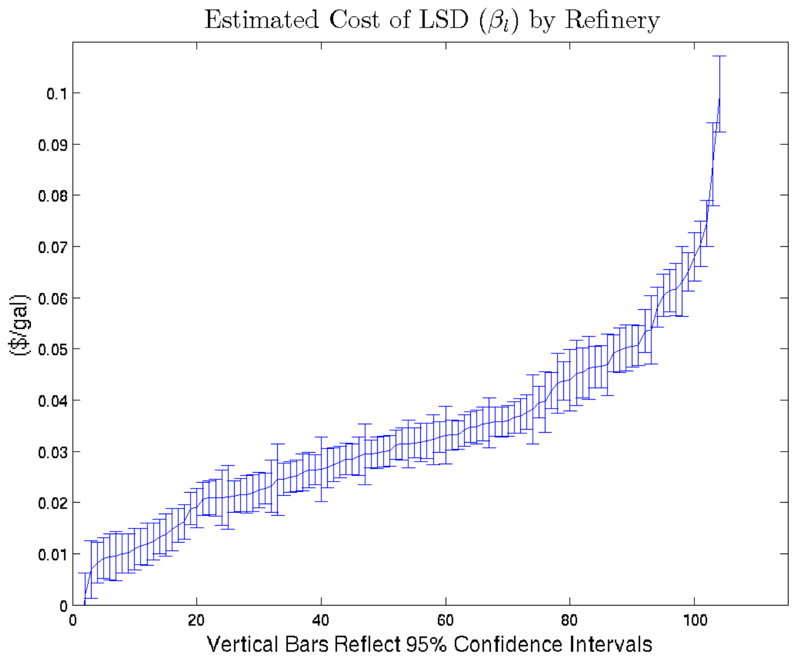
Notes: Sources: EPA (1990, 1993), NPC (1990). All costs in real (2013) cents per gallon.

Figure 1.15: RFG Cost Distribution



Turning to diesel, for a large refinery with no desulfurization capacity, LSD is estimated to cost 7.2 and 6.5 cents per gallon more than high-sulfur distillate depending on whether or not retail is included. At average desulfurization capacity levels, costs are reduced considerably to 2.9 and 3.3 cents per gallon respectively. While the intercept on desulfurization capacity is positive, the interaction with crude sulfur content and API gravity is negative, indicating that the technology reduced costs primarily by allowing refineries to process light high-sulfur crudes into low sulfur diesel. California's aromatics restriction was estimated to add an additional 2.4 cents per gallon to LSD costs in both models. Although small refineries gained valuable SO2 permits for making LSD, their costs are still estimated to be 1.1 to 1.5 cents higher than larger refineries. Presumably actual costs at these facilities are much higher. When weighted by quantity of LSD produced, the average compliance cost during the sample is estimated to be a modest 2.2 to 2.6 cents per gallon, slightly lower than estimated by EPA, and less than half what was predicted by the NPC. One reason variable costs were so low was because refineries invested heavily in desulfurization capacity after the 1990 Clean Air Act Amendments were passed. The costs of these investments are estimated in Chapter 2. Figure 1.16 plots the average LSD cost for each refinery, sorted by cost. Similar to RFG, costs vary considerably across refineries, ranging from from 0 to 10 cents per gallon, or 7 cents higher than the mean.

Figure 1.16: LSD Cost Distribution



1.5.2 Demand estimates

Demand elasticity results are presented in Table 1.8. For comparison purposes, Equation (1.2) was estimated independent of the supply side (i.e. "offline") as well as jointly. Offline OLS results are presented first, and reveal wholesale demand to be fairly inelastic, particularly for gasoline. There is reason to believe OLS price coefficients will be biased towards zero for petroleum products (Davis and Killian 2011). However, as is well documented in the literature, it is difficult to come up with valid instruments for price at the state level that have sufficient power. The third row of the table presents offline IV estimates, where regional domestic crude prices, refinery concentration, and pipeline outages were used as instruments. Only the gasoline regressions have sufficient power in the first stage, possibly due to the larger sample. For three of the four products, the IV point estimates are more elastic than OLS, but the standard errors increase dramatically, meaning that only the gasoline elasticity is statistically distinguishable from zero. The final row of the table presents the results from jointly estimating the demand and the supply sides. Demand is estimated to be much more elastic than the offline results would suggest, although the joint estimates all lie within the 95 percent confidence intervals of the IV estimates.[18] High sulfur distillate demand post regulation is estimated to be the most elastic and gasoline the least.

Table 1.8: Demand Estimates

	Gasoline	**All Distillate**	**High Sulfur Distillate**	**Low Sulfur Diesel**
Years	1986-2003	1986-1993	1994-2003	1994-2003
Uninstrumented	0.569	0.790	0.852	0.756
	(0.163)	(0.177)	(0.149)	(0.248)
Instrumented	1.119	0.425	1.157	1.268
	(0.371)	(1.326)	(2.430)	(0.910)
First Stage F-stat	25.35	4.35	11.56	8.5
Joint Estimates	1.324	2.046	4.335	2.219
	(0.032)	(0.092)	(0.208)	(0.089)

Notes: All regressions contain state-month and time dummies. Standard errors in parentheses. Offline regressions clustered at state level. Joint standard errors are robust.

For each product, the estimated wholesale price elasticities are substantially larger than more commonly reported end-use demand elasticities. For example, in a recent paper, Li, Linn, and Muehlegger (2014) estimate state-level end-use gasoline demand elasticities ranging from 0.109 to 0.365. However, if that were the relevant demand curve facing refineries, it would be difficult to reconcile such inelastic demand with observed markups. In the EIA-782 data, state-level market shares of 30 percent are not uncommon. At that level, even a relatively high end-use elasticity of 0.3 would imply Lerner markups of 100 percent above marginal costs. Yet wholesale gasoline prices are only 40 percent higher than WTI crude spot prices on average during this period. Thus, even ignoring processing costs and the fact that gasoline yields per gallon of crude are less than one, markups, at least in a quantity setting model, appear much too low to rationalize commonly reported end-use elasticities. Instead, the estimates in Table 1.8 suggest that the demand curve facing refineries, which is comprised of logistics firms and large intermediaries capable of storage, is fairly elastic.

[18] Other industrial organization papers have found supply side moments helpful in pinning down demand elasticities. For example, Berry, Levinsohn, and Pakes (1995).

1.5.3 Counterfactuals

Having recovered the cost structure of the industry and the additional costs of producing RFG and LSD, I simulated counterfactual market outcomes with these policies removed.[19] Simulation of the full model would involve solving simultaneous equilibrium in every state for every product, which was not computationally feasible. I therefore aggregated end-markets into 9 regional markets broadly reflective of pipeline patterns and refinery concentration. A map of the simulation regions is provided in Figure 1.17. Moving from state-level to region-level demand results in a minor loss of fidelity between the baseline simulated prices and observed prices in sample. The difference between the two series is less than 2 percent on average, with an R^2 of 0.983. All counterfactual results therefore compare simulated counterfactual results against simulated baseline results.[20]

Figure 1.17: Simulation Regions

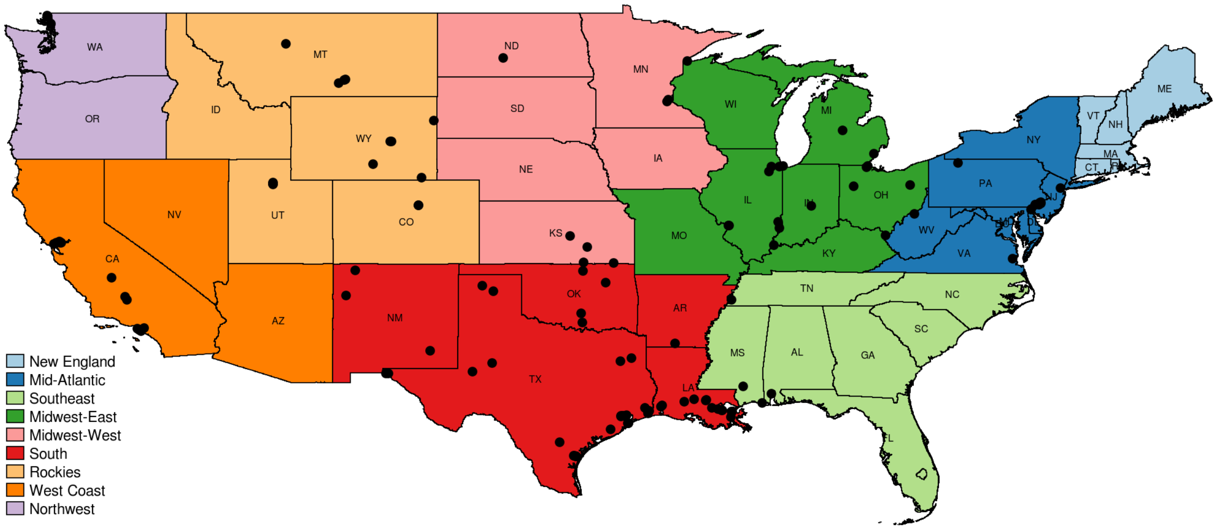


Table 1.9 presents the price results. All differences presented are the baseline outcomes under the regulation minus the outcomes with the policies removed. Gasoline prices in RFG areas are 6 to 8.2 cents per gallon higher than they would have been if those regions could have purchased conventional gasoline. However, from a national perspective, these price increases are partially offset by decreased prices in conventional gasoline markets, as refineries that found

[19] Simulations carried out in AMPL using the PATH complementarity problem solver of Steven Dirske, Michael C. Ferris and Todd Munson, available at http://pages.cs.wisc.edu/~ferris/path.html.

[20] Multiple equilibria may be possible. In order to check for this, I randomly selected 50 baseline periods and simulated equilibria for each from 10 separate starting points. In every case all 10 runs converged to the same point, suggesting that multiplicity of equilibria is not a concern in this setting.

it costly to produce RFG reallocated supply towards unregulated markets. This reallocation had the largest price effect on the West Coast, where RFG made up over 80 percent of total gasoline demand and shipment to other conventional markets was not economical. Overall, the net effect on US gasoline consumers is $14 billion over the 9 year sample.

Table 1.9: Counterfactual Results

	Gasoline Price Changes						
	Reformulated			Conventional			Average Cost
Region	**c/gal**	**%**	**Surplus ($M)**	**c/gal**	**%**	**Surplus ($M)**	**Increase (c/gal)**
New England	6.0	5.3%	-2,210	-2.0	-1.8%	85	5.7
Mid-Atlantic	6.5	5.9%	-7,772	-1.5	-1.5%	1,115	5.4
Southeast	-	-	-	-1.1	-1.1%	2,293	
Midwest - East	8.2	7.2%	-3,811	-1.4	-1.4%	2,510	5.7
Midwest - West	-	-	-	-1.2	-1.1%	720	
South	6.3	5.8%	-2,617	-1.2	-1.2%	1,914	4.5
Rocky Mountain	-	-	-	-1.5	-1.4%	590	
West Coast	7.0	5.4%	-9,241	-3.7	-3.4%	973	5.9
Northwest	-	-	-	-3.4	-3.0%	1,286	
Total			**-25,651**			**11,486**	

	Distillate Price Changes						
	Highway Diesel			Non-Highway Distillate			Average Cost
Region	**c/gal**	**%**	**Surplus ($M)**	**c/gal**	**%**	**Surplus ($M)**	**Increase (c/gal)**
New England	2.6	2.6%	-126	-0.9	-1.0%	98	2.3
Mid-Atlantic	2.3	2.4%	-855	-1.3	-1.4%	619	2.4
Southeast	2.5	2.6%	-1,164	-1.3	-1.4%	258	2.7
Midwest - East	2.5	2.6%	-1,410	-1.2	-1.4%	337	2.5
Midwest - West	2.7	2.8%	-704	-1.1	-1.2%	87	1.8
South	2.4	2.5%	-1,716	-1.4	-1.6%	478	2.7
Rocky Mountain	2.5	2.5%	-439	-2.8	-2.9%	121	4.0
West Coast	3.4	3.2%	-1,267	-5.7	-6.4%	132	4.2
Northwest	2.7	2.7%	-276	-0.7	-0.8%	44	3.3
Total			**-7,957**			**2,172**	

Notes: All numbers presented are changes relative to the counterfactual where content restrictions are removed (for example, refinery profits per gallon with RFG and LSD in place minus profits in a world with no content restrictions). The last column reports the average incremental cost of producing RFG and LSD relative to conventional gasoline and distillate at refineries serving each region.

Turning to distillate markets, price changes in highway diesel markets are pretty uniform across the country, increasing by about 2.3 to 3.4 cents per gallon. The largest prices increases were on the West Coast, where California required a more costly form of highway diesel. Similar to RFG, LSD restrictions drive down the price of high sulfur distillate, partially offsetting the total cost of the policy. In combination, the net effect on distillate consumers was $6 billion in lost consumption.

The last column reports the average increase in marginal costs at refineries supplying each RFG and LSD market. In each region, RFG market prices increase by more than costs, with average markup increases ranging from 0.3 to 2.5 cents per gallon. For LSD, costs increase by more than prices for 6 of the 9 markets, indicating that LSD reduced distillate margins on average. Table 1.10 presents the combined impact on refineries by PADD. Profits decline by 0.45 cents per gallon on average. Refineries in the Gulf Coast are least affected, having relatively low costs of producing RFG and LSD and access to the most end markets. PADD 4 refineries experience the biggest decline in profits, as these refineries are the least sophisticated and also serve the Northwest, which saw relatively larger declines in conventional gasoline prices. In sum, profits were $10 billion lower than they would have been without content regulation, an 8 percent reduction.

Table 1.10: Counterfactual Results - Refinery Profits

Padd	**Region**	**c/gal**	**($M)**	**%**
1	East Coast	-0.79	-1,646	-12.1
2	Midwest	-0.60	-3,070	-9.4
3	Gulf Coast	-0.24	-2,399	-5.1
4	Rocky Mountain	-1.48	-1,107	-24.8
5	West Coast	-0.45	-1,475	-7.4
	Total	**-0.45**	**-9,696**	**-8.2**

Notes: All numbers presented are changes relative to the counterfactual where content restrictions are removed (for example, refinery profits per gallon with RFG and LSD in place minus profits in a world with no content restrictions).

For RFG, the estimated price increases can be compared to previous estimates from the literature.[21] Brown et al. estimate RFG price effects using an unweighted average of unbranded

[21] As far as I am aware, no other studies have estimated the price impact of low sulfur diesel.

weekly city-level wholesale gasoline prices from 1994 to 1998. Their estimates range from -1.7 to 10.1 cents per gallon across cities, with a mean of 4.1.[22] The estimated price increase for RFG regions in this paper during the same period ranges from 5.7 to 7.6 cents per gallon. Although these estimates falls within the range estimated by Brown, the mean is slightly higher.[23] Using state level data and a three stage least squares approach that instruments for the size of each state's RFG market, Chakravorty, Nauges, and Thomas (2007) find that moving from zero to 100 percent RFG would increase a state's wholesale gasoline prices by 15.1 percent (16 cents per gallon on average).

1.5.4 Discussion

In Table 1.5, reduced form productivity regressions showed that, if anything, RFG and LSD were associated with increases in gasoline and distillate output per unit of capacity during the sample. However, by explicitly accounting for changes in operating incentives across time, structural estimation recovered large statistically significant marginal cost increases associated with producing these products. Simulation sheds additional light on what's going on here, revealing that refineries which found it costly to produce RFG and LSD were not only excluded from these markets, but also experienced lower margins in conventional markets.

The policy counterfactuals also highlight the importance of accounting for interactions between regulated and unregulated markets when estimating the costs of environmental regulation. Previous reduced-form studies of the Clean Air Act have implicitly calculated the gross effect of regulation on regulated versus unregulated areas or firms, leaving the net impact ambiguous. One such paper, Greenstone (2002), summarizes the importance of this distinction,

> It would be informative if the estimated regulation effects could be used to determine how much production (and employment) was shifted abroad as a result of the non-attainment designations. This would provide one measure of the national costs of these regulations. Unfortunately, such a calculation is not possible because it cannot be determined whether the lost activity in non-attainment counties moved to foreign

[22]Estimates, taken from Table 8 and average from Table 5 in Brown et al. 2008. Results in that paper were reported in nominal dollars, and are converted here to 2013 dollars to match the estimates in this paper.

[23]There are several possible reasons for this. With city level data, Brown et al. are able to pick up within-state variation that is not captured in the EIA data. To the extent that RFG areas within a state had higher prices to begin with, state-level data could overstate the price change associated with the shift to RFG. Another difference is that their price data is one step further downstream from the demand curve estimated in this paper. Their data includes resellers and marketers, as well as refineries. Finally, Brown et al. do not have quantity information, and therefore use a straight average. If sellers with low volumes also changed their prices less, then this stickiness would bias estimates of the average price effects downwards.

> countries or attainment counties. Since it is likely that the regulation effects partially reflect some shifting of manufacturing activity within the United States, they probably overstate the national loss of activity due to the non-attainment designations. Moreover, the possibility of intra-country shifting means that the regulation effects are also likely to overstate losses in non-attainment counties. The reason is that the identification strategy relies on comparisons between non-attainment and attainment counties, which leads to 'double counting' when production is moved from a non-attainment county to an attainment one.

In this paper, I find that if reformulated gasoline and low sulfur diesel restrictions were removed, approximately one third of the consumption surplus gained by consumers in those markets would be offset by decreased surplus in conventional gasoline and distillate markets.

Although the focus of this paper has been on the costs of fuel content regulation, it is important to relate them to the benefits of these programs. EPA claims that RFG Phase I and II reduced smog-causing pollution and toxics by 17 percent and 22 percent respectively compared to conventional gasoline, resulting in emissions reductions of 64,000 and 105,000 tons per year (EPA 1999). However, subsequent study has called into question whether any ozone benefits actually materialized. Auffhammer and Kellogg (2011) use detailed pollution monitor data to study changes in ozone concentrations around the programmatic and seasonal fuel restriction changes. They find no evidence that RVP regulations or federal reformulated gasoline improved ground level ozone. They do, however, find economically and statistically significant improvements in California, where stricter limits were placed on how refineries could comply with the regulation.

As far as I am aware there has been no retrospective empirical study of the benefits of the low sulfur diesel program specifically. Assuming that the pre-1990 diesel sulfur content levels would have persisted in the absence of this policy, the move to LSD represents an 80 percent reduction in sulfur emissions from diesel, which would correspond to a 1 to 2 percent reduction in national SO2 emissions during this period. It is estimated that approximately 12 percent of urban sulfur dioxide emissions are converted to particulate matter (PM) in the atmosphere. Retrospective reviews of the costs and benefits of the Clean Air Act by EPA found that benefits exceeded the costs by an order of magnitude, with most of the benefits attributed to reductions in premature mortality due to reductions in ambient PM (EPA 2011). Low sulfur diesel was implemented concurrent with several other programs targeting PM, as well as regulations on heavy duty truck engines, and a full calculation of the additional gains from LSD specifically is beyond the scope of this paper. However, the relatively modest costs of low sulfur diesel estimated here suggest that if even a small percentage of realized PM reductions during this time frame are attributable to this program, then those benefits alone would far outweigh the costs for this policy.

1.6 Conclusion

I estimate the impact of fuel content regulations imposed by the 1990 Clean Air Act Amendments on the US oil refining industry. In doing so, I account for the existence of spillovers between regulated and unregulated markets and imperfect competition. I find that reformulated gasoline increased production costs by 7 cents per gallon on average, and that low sulfur diesel increased costs by 3 cents. These costs varied significantly across refineries, resulting in gains for some refineries, particularly those able to cheaply produce RFG. However, I find that the demand curve facing refineries is significantly more elastic than would be gathered from looking at end-use consumer price responsiveness alone. As a result, operating profits were 8 percent lower from 1995 to 2003 than they would have been absent the policy[24]. These cost increases translated to consumer surplus losses of \$2.85 billion and \$884 million per year in RFG and LSD markets. However, these losses were partially offset by consumer surplus gains of \$1.28 billion and \$240 million in non-regulated markets as refineries reallocated production.

This paper has only considered reformulated gasoline and low sulfur diesel standards as they were actually implemented. However, the wide heterogeneity in refinery productivity and compliance costs estimated here suggest that, when feasible, incentive-based regulation could substantially increase cost effectiveness in this industry. Although RFG is probably not conducive to a market based approach, such as cap and trade, given the steep damages associated with local ozone concentrations, subsequent sulfur regulations and pending carbon dioxide regulations appear ideal candidates.

Finally, this paper has primarily focused on the static costs of content regulation, taking capacity investments, mergers and acquisitions as given. Yet the relatively large profit impacts estimated above are likely to have also affected investment decisions and potentially even expe-

[24]This decline in in profits only includes operating profit changes. The fixed costs associated with the low sulfur diesel program are estimated in Chapter 2

dited closures. Understanding the dynamic implications of fuel content regulation, such as its impact on long run industry concentration and point source refinery emissions, is an interesting topic for future research.

Chapter 2

Estimating the Sunk Costs of Environmental Regulation in a Large Concentrated Industry

2.1 Introduction

Environmental regulation often affects firms' fixed costs, typically by requiring that new pollution control technology be installed or by placing restrictions on the installation or expansion of other capital at firms located in polluted areas. For many environmental policies, these fixed costs are likely to be as important as, if not more important, than any variable costs. However, obtaining estimates of these sunk costs is notoriously difficult. Accounting data only includes information on actual outlays, and does not capture true opportunity costs.[1] Moreover, simply looking at costs conditional on investment may provide a biased estimate of true costs if the goal is to use those costs to study policy counterfactuals. A common reduced form approach to estimating these effects involves regressing capacity changes onto measures of regulatory stringency (Greenstone 2002, Becker 2005). However, in equilibrium contexts, even exogenous assignment of regulatory stringency can produce biased estimates, as less regulated firms' profits, and therefore investments, are affected by their competitor's investments (Chapter 1). Furthermore, these estimates cannot be translated into economically interpretable values or policy counterfactuals without imposing unmodeled structure ex-post.

A large literature in industrial organization attempts to structurally recover fixed costs by explicitly modeling market dynamics (Ackerberg, Benkard, Berry, and Pakes 2007). By directly

[1] For example, downtime while installing new capacity.

relating investment decisions to firm profits, the fixed costs of policy changes can be inferred by comparing the additional return required to generate similar investment decisions with and without the policy. While this approach generates an appealing, internally consistent analytical representation of industry dynamics, taking these models to data presents significant challenges in many settings of interest. For industries with even a modest number of players and important state variables, directly solving the model is not computationally feasible. Two-step approaches to dynamic estimation solve the computational burden, but introduce small-sample statistical problems in many settings relevant for energy and environmental regulation.

In this paper, I show how a two-period approach can be used to recover economically interpretable regulatory cost estimates in settings where fully dynamic models are not estimable. By modeling regulatory compliance as a one time decision, estimation can be carried out using moment inequalities (Pakes, Porter, Ho, and Ishii (2015)). This framework allows for the possibility of multiple equilibria and does not rely on parametric assumptions about the error distribution. More importantly, under the assumptions of a Bayes-Nash equilibrium, only unilateral deviations need to be considered, making estimation feasible for settings where the static game is complex or involves a large number of state variables.

This framework is used to estimate the sunk costs incurred by the United States oil refining industry as a result of the low sulfur diesel (LSD) program. Promulgated under the 1990 Clean Air Act Amendments, this program placed strict limits on the allowable sulfur content of distillate fuel used by highway vehicles. Sulfur can be cheaply removed from diesel by running the fuel through a desulfurization unit at the end of the refining process, and the introduction of this restriction induced a 35 percent increase in installed desulfurization capacity. Taking estimates which relate installed desulfurization capacity to variable LSD costs from Chapter 1, I estimate the sunk cost of these investments. For every refinery, I compute discounted profits under alternative investment scenarios above and below what is observed in the data, holding other refineries' decisions fixed. The intuition is that fixed costs could not have been too high, or we would not have seen the capacity investments observed; And they could not have been too low, or we have seen even more investment.

The main result is that the sunk costs of the low sulfur diesel program were between \$2.8 and \$3.3 billion. The upper bound of this range is more than 25 percent cheaper than the

National Petroleum Council (NPC) predicted the industry would have to spend on sulfur removal technology in order to comply with the low sulfur diesel program (NPC 1993).[2] These costs came in addition to the approximately $10 billion in lost operating profits estimated in Chapter 1, meaning that failing to account for sunk costs would have understated the costs of this program by 30 percent[3]. I also provide evidence that desulfurization investment would have been substantially larger if the refining industry was less concentrated, suggesting that one important channel through which imperfect competition can affect regulation costs is through underinvestment in compliance technology.

These results complement Ryan (2012)'s study of the impact of the 1990 Clean Air Act Amendments on the Portland cement industry. Using the fully dynamic two-step approach of Bajari, Benkard and Levin (2007), he finds that failing to account for the impact on fixed costs would cause one to conclude that the industry actually benefited from environmental regulation. As discussed above, this paper takes a two period approach attractive in settings where two-step models are likely to have poor small sample properties, such as the refining industry, where static profits depend critically on a large number of state variables and relatively few markets or time periods are observed. Although the timing assumed is similar to the entry and exit models models reviewed in Berry and Reiss (2007), estimating that structure within an inequality framework allows for a semiparametric approach, and yields statistically meaningful results even in industries where only a single market is observed, providing that a static model of the industry can be estimated offline. While only a single decision period is modeled here, Wollmann (2014) shows that similar methods can be used to estimate repeated games under the assumption that firms use hurdle rates when making investment decisions. Sweeney and Wollmann (2015) use simulation to evaluate the implications of this assumption for firm profits and policy outcomes in a common class of entry and exit models.

The remainder of the paper proceeds as follows. Section 2.2 provides a brief overview of the refining industry and describes the low sulfur diesel program. Section 2.3 develops a structural model of refinery investment behavior and Section 2.4 describes the estimation

[2]The National Petroleum Council is an advisory committee representing oil and natural gas industry views to the United States Secretary of Energy.

[3]The variable costs estimated in Chapter 1 include both the cost of low sulfur diesel and a the concurrent switch to reformulated gasoline. Although the switch to reformulated gasoline also involved fixed costs, a similar calculation cannot be performed because there is no specific observable capital associated with producing RFG.

procedure. Section 2.5 presents the main results of the paper, and Section 2.6 concludes with a brief discussion and plans for future work.

2.2 Background and Data

This section provides a brief overview of the US refining industry and relevant environmental regulation. In the interest of brevity, only the aspects directly relevant for estimating the fixed costs of the low sulfur diesel program are discussed in detail. For a more thorough review of the refining process and a description of the data used in estimation, see Chapter 1.

Refineries lie at the middle of the US transportation fuel supply chain (Figure 1.1). Crude oil is extracted upstream, domestically or abroad, processed at a refinery, and then shipped out via pipeline or barge to wholesale terminals, where it is distributed by truck for local consumption. Crude oil as it comes out of the ground is a mixture of different length hydrocarbons, ranging from short hydrocarbons, which roughly correspond to butane and gasoline, to long hydrocarbons, which correspond to asphalt and tar. At the most basic level, oil refining consists of separating crude into streams of differing densities using heat and a complex series of catalytic processes (Figure 1.2). The "lighter" end products, which include gasoline, diesel and propane, are typically of much higher value. So, all else equal, a refiner tries to maximize the amount of light outputs produced from a given amount of crude oil.In order to facilitate new particulate emissions standards on heavy duty diesel engines, the CAAA capped the sulfur content of on-highway diesel at 500 ppm starting in October 1993. Distillate fuel oil is a general classification of relatively heavier petroleum products used for domestic heating, industrial burners, and compression in ignition engines. Diesel fuel is distillate fuel burned in diesel engines. "On-highway" diesel fuel is diesel fuel used by trucks and passenger cars, whereas "off-highway" diesel is used in farms, construction, and marine vessels. Distillate fuel oil is thus primarily categorized by end use, rather than physical properties. This distinction is important, because the new distillate regulations imposed by the CAAA only applied to one of these categories. Home heating oil and other similar distillates were not required to meet the standard, which affected 46 percent of distillate demand, and around 8 percent of total petroleum demand at the time of enactment[4].

[4]Marion and Muehlegger (2008) provide convincing evidence that these regulations were widely enforced. Prior to 1994, they show that demand for diesel fuel and heating oil was jointly determined, as diesel consumers sought to evade taxes by purchasing untaxed distillate intended for off-highway use. Concurrent with the

Excluding these end-use categories from the sulfur requirements resulted in considerable heterogeneity in the fraction of distillate that was regulated under the program across states. Figure 2.1 presents a map with the share of each states distillate consumption that was subject to the regulation after its inception. In the northeast, substantial amounts of distillate are used for home heating oil, meaning that less than 50 percent of those markets were covered under the new regulation.[5] At the other extreme, in the southwest, over 90 percent of distillate sales after 1994 were highway diesel. Given the incomplete nature of the pipeline network and the high cost of shipping fuel via other means, this variation in the share of end-use demand distillate demand which was low-sulfur translated into heterogeneity in the extent to which different refineries were affected by the regulation. This can be shown econometrically by projecting the share of each refinery's post-1994 distillate sales that are low sulfur onto the average share of LSD post-1994 in states that each refinery was serving prior to 1990, when the regulation was announced. Doing this reveals that a 10 percent increase in LSD share in states served by a refinery prior to the regulation leads to a 13 percent increase in the fraction of distillate that refinery converts into LSD (Table 2.1).

Table 2.1: Geographic Determinants of RFG and LSD

	Post 94 % LSD	Δ Desulf Cap 90-96
% States Served Pre 1990	1.341***	0.900***
	(0.421)	(0.235)
N	122	121
F	10.14	14.70

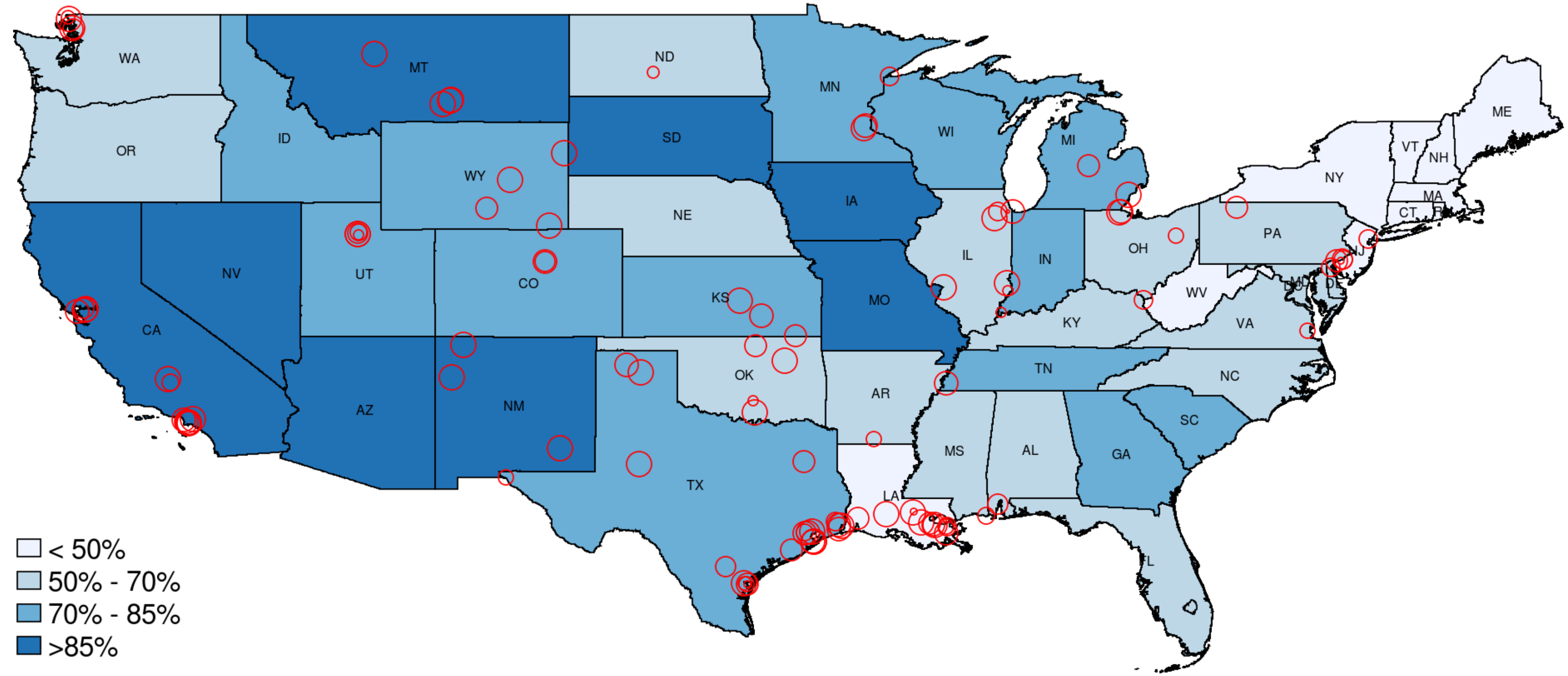
Notes: The independent variable in each regression is the share of each refinery's pre-1990 markets which subsequently became regulated post-1994. The dependent variable in the first model is the share of each refinery's distillate production which was low sulfur in the post period. The dependent variable in second column is the change in desulfurization capacity per unit of distillation capacity.

Producing low sulfur diesel was a significant achievement for the industry, as the national average sulfur content of distillate at the time was 3000 ppm (Lidderdale 1993). The primary method of compliance involved sending distillate through a desulfurization unit after distillation. In Chapter 1, desulfurization capacity was estimated to significantly reduce the incremental

introduction of low sulfur diesel, the government mandated that non-highway distillate be marked with a dye to prevent illegal sales.

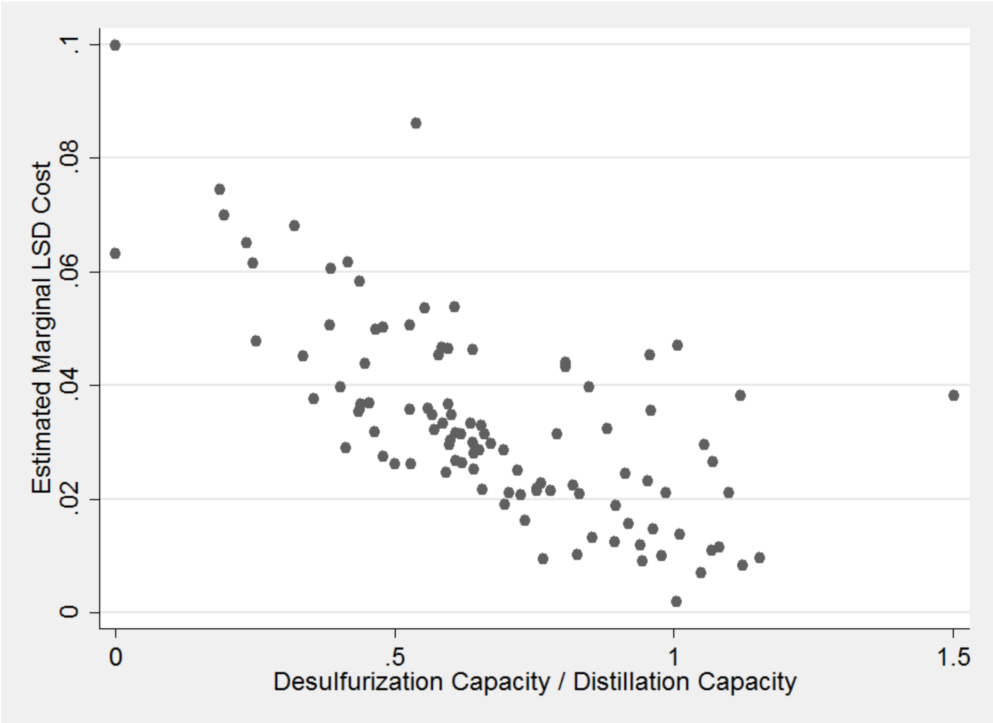
[5] In other parts of the country, off-highway distillate is mainly used in farming equipment or marine vessels.

Figure 2.1: Share of Post-94 Distillate Sales that are Low-Sulfur



NOTE: Circles proportional to the average fraction of distillate that is low-sulfur for each refinery variable cost of converting conventional distillate in to LSD (Figure 2.2). All else equal, a 50 percent increase in desulfurization leads to 1.5 cent decrease in incremental costs per gallon.

Figure 2.2: Estimated Marginal LSD Cost vs. Desulfurization Capacity

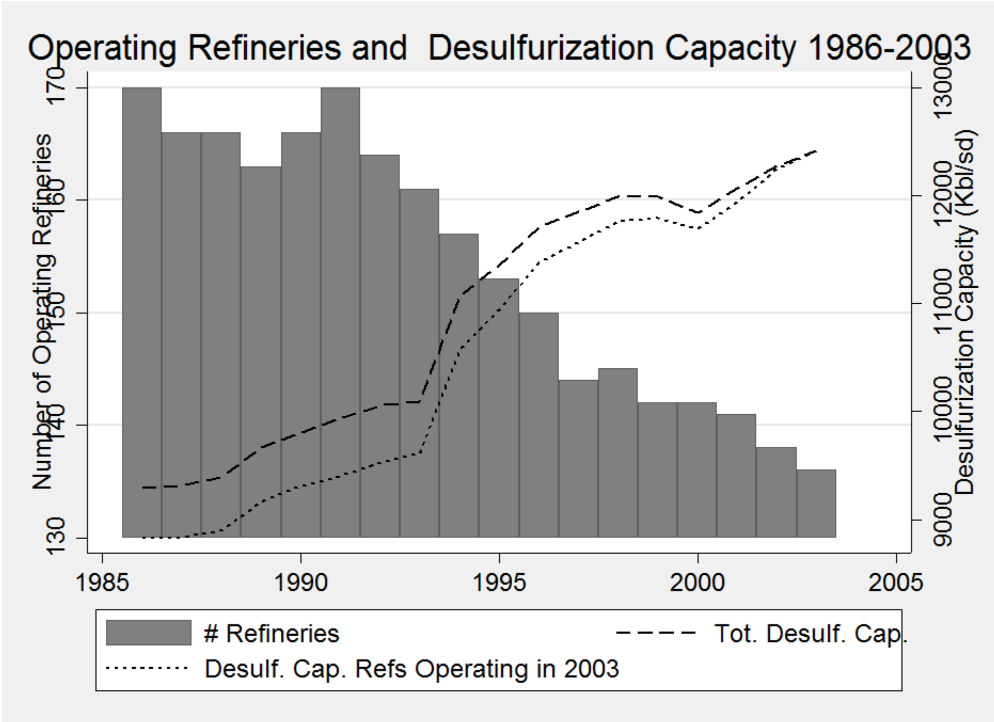


In preparation for the switch to low sulfur diesel, refineries invested heavily in desulfurization. Figure 2.3 shows that total desulfurization capacity increased by almost 40 percent from 1990 to 2003, the end of the sample, with almost all of those additions coming prior to 1997.[6] For comparison purposes, the trend in distillation capacity is reported for the same period in panel (b). It's clear that this increase was not simply the result of a general expansion in refining capacity. A projection of the ratio of these capacity changes onto the measure of refinery exposure from the last paragraph finds that these investments were significantly determined by the relative size of the LSD market facing each refinery (Table 2.1). A 10 percent increase in LSD share in states served by a refinery prior to the regulation was associated with a 9 percent increase in installed desulfurization capacity. While this regression provides strong evidence that the low sulfur diesel program resulted in large investments, it is not clear what this coefficient translates into in terms of actually costs. In order to do this, the next section develops

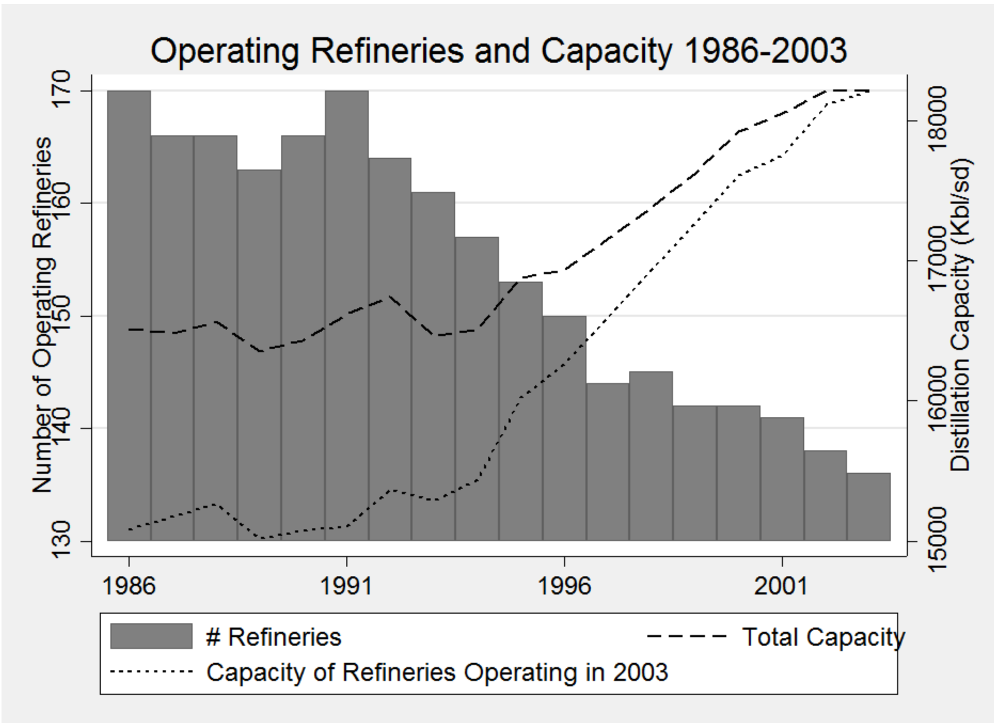
[6] In this paper, I focus on reformulated gasoline (RFG) and low sulfur diesel (LSD) from 1994 to 2003. In 1999, new sulfur limits were announced for gasoline and highway diesel, which phased in starting in 2004 and 2006 respectively.

a model which directly relates changes in profits from additional desulfurization to observed investment decisions.

Figure 2.3: Capacity Trends by Exit Status



(a) Desulfurization

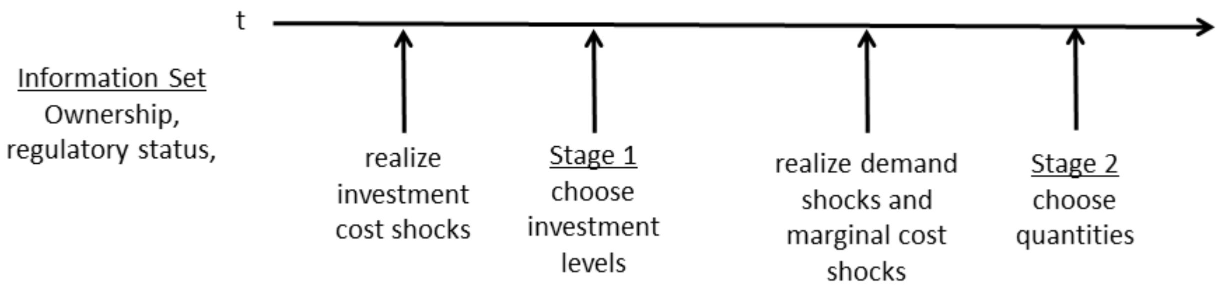


(b) Distillation Capacity

2.3 Model

In this section I develop a model of refinery investment. Figure 2.4 presents a timeline of the game. At the start of each period, firms have an information set ($\mathcal{J}_i$) which includes knowledge of the ownership status and installed technology of every refinery in the industry and information about the state of environmental regulation. They then receive a shock to the cost of making changes to their capital stock and chose investment. After all investments come to fruition, firms realize operating cost and demand shocks and compete simultaneously in quantities for each product in each state.

Figure 2.4: Timing of the Game



Stage game profits for firm f owning refineries $i \in I_f$ in month t are taken from Chapter 1.

$$\Pi_{ft} = \sum_{i \in I_f} \sum_j \sum_m (P(Q_{jmt}, \epsilon^D_{jmt}) - \tau_{im}) q_{ijm} - Cost(\mathbf{q_i}, X_{it}, K_{it}, \epsilon^S_{it}) = \Pi_f(\mathbf{s_t}, \epsilon_\mathbf{t})$$

Where K_{it} is a vector of installed capacity, X_{it} is a vector of other variables affecting production costs, and $\mathbf{s_t}$ is a vector of all refineries' cost function inputs at time t. Demand is assumed to be constant elasticity, and depends on both observables in the firm's information set and unobserved demand shocks realized after investments are made. Months t are grouped into larger time blocks $b = 1...B$, each containing T months. Define the discounted sum of profits for an entire time block as

$$\Pi_{fb}(\mathbf{s_b}, \epsilon_\mathbf{b}) = \sum_{t=1}^{T} \beta^t \Pi_{ft}(\mathbf{s_t}, \epsilon_\mathbf{t})$$

Let $d_{ib} \in D$ be a decision by refinery i which governs the change in capacity from time block $b-1$ to b. These decisions come at sunk costs $SC(d_{ib}, \nu_{ib})$ where ν_{ib} is a private investment cost shock. Given the current state vector and investment cost shock, firms maximize the expected sum of discounted profits

$$E\left[\sum_{b=0}^{B} \beta^{(b-1)T} (\Pi_{fb}(\mathbf{s_b}, \epsilon_\mathbf{b}) - SC(d_{ib}, \nu_{ib}))\, |\mathbf{s_0}, \nu_\mathbf{0}\right]$$

Define strategy $\sigma_i(\mathbf{s}, \nu_i)$ which maps (possibly a history of) states and cost shocks into decisions, and define $V(\mathbf{s}, \nu_i|\sigma_i, \sigma_{-\mathbf{i}})$ as the expected discounted profits for firm i under the prevailing strategies of all players. The workhorse framework for handling such dynamic games is to assume Markov perfect equilibrium[7]. In this setup, strategies are a function of the current state and shock only, and provide the maximum expected discounted profits for each firm given the strategies of its competitors. Under these assumptions, the expected profit function firms maximize can be written in the familiar recursive form.

$$\begin{aligned} V_i(s, \nu_i|\sigma) \quad &= \quad -SC(\sigma(s, \nu_i), \nu_i) + E_{\nu_{-i}}\Big[(\Pi_{fb}(\sigma(\mathbf{s}, \nu), \epsilon_\mathbf{b}) \\ &\quad +\beta \int V_i(\mathbf{s}', \nu'|\sigma) dG(\nu_i'|\mathbf{s}') dP(\mathbf{s}'|\sigma(\mathbf{s}, \nu), \mathbf{s})\Big] \end{aligned}$$

As was discussed in Section 1, estimating a full Markov perfect equilibrium of the refining industry is not feasible. Under any reasonable specification, $\mathbf{s}$ will be large and complicated enough that directly iterating on V would involve storing and taking the expectation over billions of states. One the other hand, two-step approaches to estimation are also problematic, as they rely heavily on the observed empirical transition matrix across states in the data. In the refining industry, competitive regions overlap in such a way that s will have to contain multiple dimensions of information on all refineries in the US. Given the small number of refineries, such approaches are likely to have poor small sample properties, due to the fact that transitions

[7]For a thorough review of methods for estimating dynamic games, see Ackerberg, Benkard, Berry, and Pakes (2007)

between any two of these states are observed with very low probability.

In order to circumvent these issues, I consider desulfurization investment decisions independently from all other dynamic decisions, and model changes induced by the low sulfur diesel program as a one time decision[8]. After learning about the new low sulfur diesel requirement, refineries each receive a private shock to the cost of investing in desulfurization technology and then simultaneously make investments. They then receive static profits resulting from these choices each month thereafter.

Under these one-shot game assumptions, where $b = 1$ and $T = \infty$, the value function above reduces to

$$V_i(s, \nu_i|\sigma) = -SC(\sigma(s, \nu_i), \nu_i) + E_{\nu_{-i}} \left[(\Pi_{fb}(\sigma(\mathbf{s}, \nu), \epsilon_{\mathbf{b}}) \right] \tag{2.1}$$

2.4 Estimation

The model is estimated using the data described in Chapter 1. The data includes input and output quantities for every refinery in the United States by month, and detailed information on the capital installed at each refinery at the start of the year. The sample begins in 1990 when the Clean Act Amendments are passed. I then allow sufficient time to pass for these investments to come on line, which, looking at Figure (2.3a), appears to have happened by 1997. Additional sulfur regulations were announced in 1999, and these investments begin to come on line in 2004. I therefore take the period from 1997 to 2003 as the relevant sample. The sample is restricted to 109 large, non-specialty refineries operating post 1997.[9]

Estimation proceeds using the moment inequality approach introduced by Pakes, Porter, Ho, and Ishii (2015) (henceforth, PPHI). This approach does not require knowing the distribution of ν, and allows for the possibility of multiple equilibria. The general intuition is a combination of rational expectations and Nash equilibrium: in expectation, observed decisions should be more profitable than feasible alternatives conditional on the realized choices of all other agents. As

[8] Another alternative is to model dynamic decisions as repeated one shot games. See Wollmann (2014) and Sweeney and Wollmann (2015).

[9] This excludes 11 refineries which exited between 1990 and 1997 from the sample used in Chapter 1, as they do not operate in the post period. Although it is unlikely, if these exits were precipitated by large $\nu_2$draws, then this could bias the θ estimated here downward. I also exclude two sets of refineries that merge during the sample. These four refineries are excluded from estimation, as the acquiring refineries' desulfurization capacity increases, but these increases presumably do not come at cost SC.

such, estimation only requires computing profits for a small subset of the state space, making it tractable for even complicated stage games with many states.

Placing the model within the framework of PPHI, let $d_i = k_{i1} - k_{i0}$, where k_{i1} and $k_{i1}$are the amounts of desulfurization capacity installed at refinery i before and after the policy is announced. The sunk costs of increasing desulfurization capacity are assumed to be linear in d,

$$SC(d_i, \nu_i) = d_i(\theta + \nu_{2i})$$

ν_2 is an i.i.d. mean zero cost shock which was known to the refinery when it made its decision. This gives refineries a payoff function

$$\pi(d_i, \mathbf{d_{-i}}, \mathbf{z_i}) = R(d_i, \mathbf{d_{-i}}, \mathbf{z_i}) - d_i(\theta + \nu_{2i}) + \nu_{1id_i}$$

Where R is the discounted profits from the post-period, $R(d_i, \mathbf{d_{-i}}, \mathbf{z_i^o}) = \Pi_{fb}(\mathbf{s_b}, \epsilon_\mathbf{b})$. $\nu_{1,i,d}$ are mean zero expectation and measurement errors in the refinery profit function. Differences in the payoff function for alternative investment choices are denoted

$$\Delta\pi(d, d', \mathbf{d_{-i}}, \mathbf{z_i}) = \Delta R(d, d', \mathbf{d_{-i}}, \mathbf{z_i}) + (d' - d)(\theta + \nu_{2i}) + \nu_{1idd'}$$

The assumptions of Bayes-Nash equilibrium and simultaneous decisions yield the following inequality

$$E[\Delta\pi(d, d', \mathbf{d_{-i}}, \mathbf{z_i})] \geq 0 \quad \forall d \in D \tag{2.2}$$

This equation implies that observed choices were best responses in expectation. If ν_2 was observed, we could take this equation to data. However, ν_2 is unobserved to the econometrician, and generates selection in observed decisions. PPHI show that θ can be consistently estimated if we have a suitable non-negative weighting function, or instrument, $h^i(d'; \mathbf{d_i}, \mathcal{J}_i)$ such that

$$E[\sum_{i=1}^{n} \sum_{d' \in D} h^i(d'; \mathbf{d_i}, \mathcal{J}_i)(d' - d)\nu_{2i}] \leq 0 \tag{2.3}$$

$$E[\sum_{i=1}^{n}\sum_{d'\in D} h^i(d';\mathbf{d_i},\mathcal{J}_i)\nu_{1i}] \geq 0 \tag{2.4}$$

Combining assumptions and replacing $\pi(d_i, \mathbf{d_{-i}}, \mathbf{z_i})$ with observable approximation $r(d_i, \mathbf{d_{-i}}, \mathbf{z_i},\theta_0)$ yields the following population condition

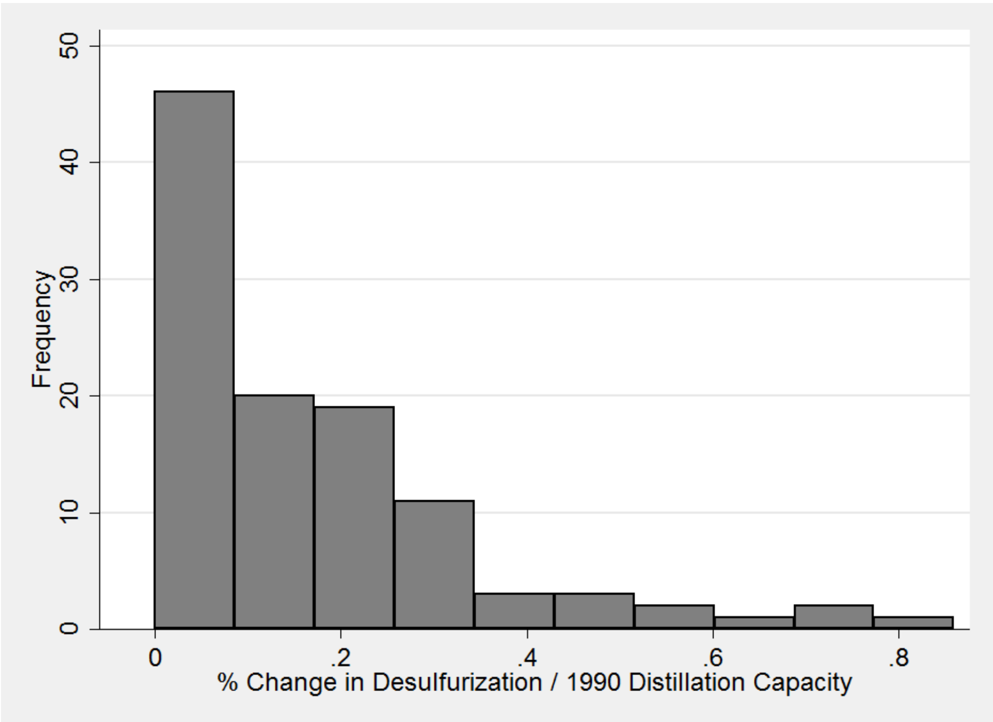
$$E[\sum_{i=1}^{n}\sum_{d'\in D} h^i(d';\mathbf{d_i},\mathcal{J}_i)\Delta r(\mathbf{d_i}, d', \mathbf{d_{-i}}, \mathbf{z_i^o}, \theta_0)] \geq 0 \tag{2.5}$$

This inequality is based entirely on observables and, given suitable choices of d, can be taken to data. Investment in desulfurization capacity is a continuous version of the ordered choice problem discussed in PPHI. Choosing d' above and below the observed d_i for each refinery generates moments which bound θ from above and below if Π is concave in k. The ordered choice problem in PPHI involves a discrete decision d adapted from Ishii (2004). In that case, $(d' - d)$ equals plus or minus 1 for all i, which means (2.3) is satisfied for any h^i uncorrelated with ν_2. In the case of refineries, which differ considerably in size, choosing a fixed deviation $\Delta d = |d' - d|$ for all refineries would be problematic, as any feasible Δd would either trivially affect large refineries or significantly affect small ones. As an alternative, I choose Δd_i to be a fixed percentage of each refinery's 1990 distillation capacity, $\Delta d_i = .01 * Cap_{90}$. Thus, satisfying (2.3) involves assuming that ν_2 is uncorrelated with 1990 capacity levels. Operationally, I implement this by taking a set of instruments $g(x_i)$, assumed to be uncorrelated with ν_{2i}, and constructing $h^i(d';\mathbf{d_i},\mathcal{J}_i) = g(x_i)/Cap_{90i}$.

Using the simulation model described in Chapter 1, I calculate $\Delta R(d_i', \mathbf{d_{-i}}, \mathbf{z_i})$ for positive and negative Δd_i deviations, for all 109 refineries operating post 1997. Calculating ΔR is complicated by the fact that the simulation model only goes through 2003, while the one-shot game assumption above implied that R should include $\sum_{t=1}^{\infty}\beta^t\Pi_{ft}(\mathbf{s_t}, \epsilon_\mathbf{t})$. I approximate for the full continuation value by assuming refineries get their 2003 stage profits in perpetuity. $\beta = 1/1.07$ is assumed, which is the upper discount rate used by government agencies when conducting regulatory impact analyses. In the results section, I conduct sensitivities to this assumption.

A separate truncation issue arises from the fact that only 87 of the 109 refineries in the sample invest in desulfurization during this period. Figure 2.5 provides a histogram of the observed changes in desulfurization as a fraction of distillation capacity. For 22 refineries, the deviation $-\Delta d_i = -.01 * Cap_{90}$ is not feasible. If these refineries did not move because they received high ν_2 draws, then excluding them from (2.5) could cause the estimated upper bounds on θ to be biased downward. I address this issue by implementing the truncation correction suggested by PPHI. Under the assumption that ν_2 is symmetrically distributed, they suggest and an $h()$ which bounds the missing $\nu_2's$ using the rank ordering of upward deviation values, for which all deviations are feasible.[10] It is worth noting that this correction could not be implemented if more than half of the observations were on the lower bound. This serves as additional motivation for the one-shot approach taken in this paper, as while over 80 percent of refineries make desulfurization investments between 1990 and 1997, only 30 percent move in a given year.

Figure 2.5: Observed percent Changes in Desulfurization Capacity



Estimation proceeds by finding the set of θ which minimize the following criterion function

$$Q(\theta) = || \sum_k (m_k(W_i, \theta)_-) ||$$

[10] The interested reader is referred to PPHI for further details on this correction.

where $m_k(W_i, \theta)$ denotes the moment formed when using instrument h_k and data W on the left hand side of equation (2.5). $(\cdot)_- = min\{\cdot, 0\}$, as equation (2.5) only implies that the moments should be positive, and does not place any weight on the constraints being "more" satisfied.[11]

Inference is based on methods proposed by Andrews and Soares (2010), as described in an earlier version of PPHI. The general idea is to construct a test by bootstrapping the criterion function Q for a sufficiently wide range of guesses $\{\theta_g\}$ under the null that $\theta_g = \theta_0$[12]. For a given guess of θ_g, I take 1000 draws from a normal distribution with mean zero and variance equal to the variance of the sample moments. I then calculate the criterion function for each draw in order to approximate its distribution under the null. θ_g is accepted if $Q(\theta_g)$ is below the 95th percentile of this estimated distribution. The minimum and maximum of this set of accepted θ_g's are then analogous to the 95percent confidence intervals commonly reported in empirical work.

2.5 Results

Table 2.2 reports the main results of the paper. θ is denominated in dollars per barrel of stream day capacity. In order to provide context for these estimates, the last two columns convert the estimated values into the implied total compliance costs of observed investments. Ex-ante, the National Petroleum Council predicted that the refining industry would have to spend $4.34 billion investing in sulfur removal technology in order to comply with the low sulfur diesel program (NPC 1993). All inequality estimates are set identified, meaning that multiple values of θ resulted in a criterion function equal to zero.

[11] Although this criterion function will provide an unbiased estimate of θ, in empirical work it is common to weight moments m by the inverse of their variance. However, when $\hat{\theta}$ is a set (rather than a singleton), as in this paper's results below, this weighting is superfluous.

[12] As θ is just a single parameter, I set the range of $\{\theta_g\}$ 50 percent below and above the bounds of the estimated set, and check that the CI estimated is not on the corner in either direction.

Table 2.2: Desulfurization Cost Estimates

	θ	95% CI for θ		Implied Cost ($B)	
	($/bsd)	LB	UB		
1. g(x) ≡ 1 , d > 0 for u.b.	[1274, 1389]	1007	1697	2.79	3.05
2. g(x) ≡ 1 , d ≥ 0	[1274, 1511]	1027	2097	2.79	3.31
g(x) = (1, % LSD 1990, Latitude)					
3. d > 0 for u.b.	[1274, 1363]	1007	1627	2.79	2.99
4. d ≥ 0	[1274, 1489]	1057	1907	2.79	3.27
Refinery profits only					
5. g(x) ≡ 1 , d > 0 for u.b.	[1653, 1783]	1297	2167	3.63	3.91
6. g(x) ≡ 1 , d ≥ 0	[1653, 1937]	1367	2557	3.63	4.25
Alternative discount rates					
7. β=1/1.05 , g(x) ≡ 1 , d ≥ 0	[2001, 2381]	1591	3241	4.39	5.22
8. β=1/1.10 , g(x) ≡ 1 , d ≥ 0	[756, 893]	618	1198	1.66	1.96
First Order Conditions					
9. g(x) = 1	1342	1133	1551	2.94	
10. g(x) = (1, % LSD 1990, Latitude)	1323	1119	1527	2.90	

Row 1 reports the results where $g(x)$ is a constant. Although θ is not a singleton, the estimated bounds on the true value are quite narrow, with the upper bound only 9 percent higher than the lower bound. However, as discussed above, the sample for the upper bound is truncated by the fact that $-\Delta d$ deviations were not feasible for every refinery. Row 2 reports the results when the PPHI truncation correction is imposed. Although this doubles the range of the estimated set, the bounds are still quite tight, with the upper bound now 19 percent higher than the lower bound. Under these estimates, the implied costs of the desulfurization investments made after the announcement of the low sulfur diesel program are between $2.8 and $3.1 billion.

Rows 3 and 4 repeat the same exercise using additional instruments. As was discussed above, refineries were differentially affected by the switch to low sulfur diesel due to geographic differences in the demand for heating oil. Given this, I construct an instrument equal to the average share of low sulfur diesel post-1994 in states that each refinery was serving prior to 1990, when the regulations were announced. Table 2.1 shows that this measure is strongly correlated with the post-1994 production of low sulfur diesel. As the main alternative use for distillate fuel is home heating oil, I also use the latitude of each refinery as an instrument, which should be correlated with heating oil demand through climate. Interestingly, the estimated lower bound on θ is unaffected by the addition of these instruments, indicating that the moment with $g(x) = 1$ bounds in this direction. The upper bounds are tightened modestly, reducing estimated upper bound on investment costs by \$60 million.

In the main specification, $\Delta r(\mathbf{d_i}, d', \mathbf{d_{-i}}, \mathbf{z_i^o}, \theta_0)$ incorporates the change operating profits across all refineries owned by the firm that owns refinery i. Rows 5 and 6 repeat the exercise from rows 1 and 2 but assume that firms only consider the change in profits for the investing refinery when making investment decisions. Under this assumption the estimated bounds on θ are about 30percent higher than before. This is not surprising, as some of the increased low sulfur diesel sales refineries gain from increasing their desulfurization capacity are cannibalized from other refineries owned by the same parent firm. Although this exercise cannot shed light on which assumption is more accurate empirically, the exercise suggests that the presence of multi-refinery firms could have significantly reduced desulfurization investment, and therefore increased LSD prices, if firms behaved rationally.

Rows 7 and 8 perform a sensitivity analysis on the assumed discount rate for the truncation corrected results. Discount rates of 5percent and 10percent were considered instead of the 7percent rate assumed in the main specification. The former is approximately equal to the average rate of return reported by refineries in the EIA's Financial Reporting Service during this time period, while the latter is assumed by the National Petroleum Council when it estimates the cost of environmental regulation on the refining industry. The results are quite sensitive to this choice, with the estimated cost under a 5percent discount rate ranging from \$4.4 to \$5.2 billion. Somewhat ironically, using the higher rate assumed by the NPC results in very model implied costs, equal to less than half of the NPC's estimated capital costs for the program.

The final two rows of the table present results from an alternative estimation approach suggested by PPHI and based on Hansen and Singleton (1982). An moment condition is derived from the first order conditions of equation (2.1). With continuous controls, the firm's optimization program suggests minimizing the following criterion function

$$||\frac{1}{n}\sum_i \{d_i > 0\} \left(\frac{\partial R(d, d_{-i}, z_i^o)}{\partial d}|_{d=d_i} - \theta \right) \times h(x_i)||$$

$\frac{\partial R(d,d_{-i},z_i^o)}{\partial d}$ is approximated using the average of ΔR across positive and negative deviations for every refinery, and θ is estimated using the generalized method of moments. Row 9 reports the results using $g(x)= 1$. The point estimate of θ lies well within the range estimated using inequalities. Row 10 repeats the exercise using the additional instruments, and, as above, finds that this only marginally alters the estimated costs. In sum, this exercise seems to indicate that these two approaches are pretty similar in this setting, which is perhaps not surprising for a continuous control with concave profits in that dimension.

2.6 Conclusion

This paper demonstrates the way in which a rich static model of firm profits can be combined with assumptions on the timing of firm decisions to recover estimates of sunk costs imposed by environmental regulation. Using a moment inequality approach, I am able to estimate costs for a setting in which other approaches would not be feasible. I find that complying with the imposition of strict sulfur limits on diesel fuel involved \$2.8 - \$3.3 billion worth of investment in desulfurization capacity at US oil refineries. Failing to incorporate these sunk costs into an evaluation of the program would have understated the costs to the refining industry by 30 percent. While the estimated costs are quite large, they are significantly lower than the industry predicted when the policy was announced.

This paper also contributes to a growing empirical literature highlighting the importance of market structure in determining the realized costs of environmental regulation (Millimet et al. 2009). I find that the presence of multi-refinery firms significantly reduces the returns to investing in desulfurization technology compared to a world where all refineries operate independently. These results complement results in Chaper 1 showing that refineries operate at lower utilization rates when the owning firm's share of capacity in a region increases. In future work, I plan to structurally model the decision to merge in order to estimate the impact of market power on other capital investments in this important industry.

Finally, in this paper no counterfactuals were computed. The reason is that the goal of the analysis was to estimate the cost of a technology which would have had little operational benefit absent the policy. While there are many other similar settings, or situations where the per unit sunk costs estimates are of primary interest in their own right, other policies affecting the refining industry do suggest interesting counterfactuals. In particular, the New Source Review Program mandated that refineries located in heavily polluted areas undergo additional scrutiny before making capital investments. In future work I plan to estimate the cost of these restrictions using the repeated two-period hurdle rate setup of Wollmann (2014).

Chapter 3

The Role of Sales Agents in Information Disclosure[1]

3.1 Introduction

Consumers often learn about new or higher-quality products from firms, and in theory, information problems can be ameliorated as sellers of relatively high-quality products inform consumers of their beneficial attributes.[2] In many cases, however, barriers to information transmission can cause a failure of "unraveling," and imperfect information could reduce demand for high-quality products. In these situations, regulators may be able to increase welfare by mandating or otherwise inducing information disclosure, subsidizing quality, or even setting minimum quality standards. Before intervening, however, a regulator might want additional information: How successful are firms at providing information? How well-informed are consumers about new products and their attributes?

Every year, Americans purchase \$361 billion in energy-using durable goods such as cars and air conditioners and spend \$570 billion on energy for those goods (BLS 2014). Regulators intervene in durable goods markets by mandating energy use information disclosure and by encouraging additional marketing of energy efficiency through initiatives such as the Energy Star Retail Partner program. Imperfect information is also commonly used to justify extensive subsidies for energy efficient goods, as well as minimum energy efficiency standards.[3] Despite

[1] Co-authored with Hunt Allcott

[2] See Grossman (1981), Milgrom (1981), and Viscusi (1978).

[3] There are many examples. The American Council for an Energy Efficient Economy argues that minimum efficiency standards are merited for several reasons, including "rush purchases when an existing appliance

the importance of imperfect information in the energy policy debate, however, there is limited evidence on how energy cost information disclosure affects demand for durable goods. Dranove and Jin's (2010) review of the information disclosure literature discussed zero studies related to energy efficiency, although we discuss below how this literature has recently received more attention.

We study water heaters, which are interesting and important precisely because they are so mundane. Consumers rarely think about their water heater until it breaks unexpectedly, at which point they want to replace it quickly, with limited time for search and information acquisition. Consumers rarely discuss water heaters with their friends, meaning that retailers play a pivotal role in guiding purchases. At average product lifetimes and usage rates, purchasing an energy efficient "Energy Star" natural gas water heater instead of a standard model is an investment with 13-18 percent return - and this is before the generous subsidies offered by many local utilities. Despite this, the Energy Star market share is only about 3 percent at the retailer we study. These choices are expensive: water heating is the second largest home energy use in the U.S. (DOE 2009), consuming about \$300 in energy annually per household, or about \$29 billion per year nationwide.

Motivated by these issues, we carried out a natural field experiment with a large nationwide retailer that sells water heaters and many other goods. We worked at the Retailer's call center, which sells about 45,000 water heaters each year. More than 20,000 callers were randomly assigned between treatments in which sales agents were instructed to provide energy cost savings information and/or offer customer rebates for Energy Star models. We also offered \$25 sales incentives for agents who sold Energy Star on randomly-selected calls, and we crossed these incentives with the customer rebates.

A crucial feature of our experiment is that the seller's interactions with customers are intermediated by sales agents. This is not uncommon: consumers learn about life insurance, mutual funds, and many consumer goods at least partially through agents. In our setting, sales agent behavior is important for two reasons. First, it directly determines the Retailer's ability

breaks down, providing no time to comparison shop" (Nadel 2011). The Regulatory Impact Analysis for the increase in the Corporate Average Fuel Economy standard for 2012 to 2016 argues that even without counting the value of externality reductions, the regulation increases consumer welfare, perhaps because consumers are not correctly informed about the value of fuel economy (NHTSA 2010, page 2). The Regulatory Impact Statement for Australia's ban on energy inefficient lightbulbs argues that "information failures" help to justify that policy (DEWHA 2008, page vii).

to market Energy Star products: if sales agents don't provide information on a call, callers will likely remain uninformed. Second, in equilibrium, it is indirectly informative about consumers' responsiveness to information: given that information disclosure takes time and focus away from other sales tasks, if consumers are not interested in information, agents will not provide it. To document sales agent behavior, our research team independently audited more than 2,000 phone calls, quantifying the interactions between agents and consumers. Our ability to observe agent behavior, instead of simply the equilibrium outcome of attempting to disclose information, is one feature that distinguishes this paper from previous work.

There are several reasons to expect that the treatments could substantially increase Energy Star market share. Because consumers are thought to be poorly informed about water heater features, they often accept sales agents' recommendations about what model to purchase. The $100 customer rebate increases the average consumer's return on investment in an Energy Star model to 28-37 percent, and when combined with additional subsidies available from many local utilities, our $100 rebate brings the incremental purchase price of the Energy Star model close to zero. The $25 sales incentives are equal to two times agents' fixed hourly wage and are ten times larger than their usual sales incentive.

Against this backdrop, our results are surprising. Our audits show that agents comply with delivering the information and rebates on only about one-fifth of calls. Information has zero statistical effect on demand, and confidence intervals rule out that demand increases by more than 4.9 percentage points on calls when the information is delivered and the consumer is considering a substitutable model. While this bound is large relative to the baseline Energy Star market share, it suggests that the market share would still be very low even if agents informed all consumers. The $100 customer rebates do increase Energy Star purchases, however, and the combination of a $25 sales incentive and $100 customer rebate appears to have particularly strong complementary effects. This last result suggests one potential policy implication: when addressing market failures that might distort energy-using durable good purchases, policymakers may wish to consider also incentivizing sales agents instead of only subsidizing consumers' purchases.

We show that agents preferentially market Energy Star to consumers with higher latent demand for Energy Star. Furthermore, on calls with the $25 sales incentive but no experimental customer rebate and no explicit direction to deliver an informational script, agents exert very little effort to sell Energy Star models. Along with the small information effect, these results suggest that agents' non-compliance is better described as "strategic" instead of "shirking": agents don't inform consumers about Energy Star because they know that at the retailer's base price, most consumers are not interested in the product once informed.

There are two potential explanations for why the Retailer's attempts at information disclosure did not increase Energy Star demand. First, consumers may tend to be unaware of Energy Star and underestimate its benefits, but sales agents may not be able to address this because their "disclosure technology" is limited: they work in time-constrained sales interactions and may have a perceived lack of credibility when promoting a higher-priced model. Second, consumers might already be relatively well-informed about Energy Star availability and benefits, and most choose not to buy because they don't think that the reduced energy use is worth the incremental upfront cost. These two explanations have very different implications for whether regulators should intervene to provide information or otherwise encourage energy efficiency.

We carried out an extensive set of customer follow-up surveys to shed light on these two explanations. On the one hand, there is evidence that consumers are confused: even when "Energy Star" is precisely defined, 52 percent of consumers report believing that they had bought an official Energy Star model, while only 2.1 percent actually had. Of the consumers who thought they had not bought Energy Star, 15 percent reported that this was because they were not aware that there was an Energy Star option. On the other hand, the great majority of consumers were aware of Energy Star, and their foremost reason for not purchasing was that the price was too high. Furthermore, while there is wide dispersion of beliefs, the average consumer actually overestimates the potential energy cost savings from Energy Star. While the survey results are not as conclusive as the experimental results, they at least suggest that lack of awareness and cost savings information are not the primary barriers to energy efficiency in this context.

We also include a simple theoretical model which helps to motivate the experiment and interpret results. The model is a two-firm version of Grossman and Shapiro's (1984) analysis of informational advertising in a Hotelling spatial model. We extend this framework to include two goods, a base good and a "high-quality" (energy efficient) good, and nest within the firm a set of optimizing sales agents who disclose information at convex cost. The model highlights the importance of equilibrium interactions between the firm's management, sales agents, and consumers: because information provision is costly, sales agents will not provide information if consumers are unresponsive, and consumers will be unresponsive when the firm sets a relatively high price for the high-quality good. When the firm lowers the price of the high-quality good, however, consumers become more interested in the product, and agents exert more effort in marketing it. The model also helps to clarify the policy relevance of the experiment by formalizing how minimum quality (energy efficiency) standards are more likely to increase welfare if the high-quality good is more beneficial and if experimental results show that it is difficult for firms to inform consumers of this. Furthermore, the model formally considers government-provided sales agent incentives, showing that they can increase welfare if firms under-provide information relative to the social optimum and firms are not able to "undo" the government-provided incentive in equilibrium.

The remainder of this section discusses related literature. Section 2 provides an overview of the water heater market, and Section 3 presents the theoretical model. Section 4 details the experimental design and data. Section 5 presents the empirical results, and Section 6 concludes.

3.1.1 Related Literature

Our study is broadly connected to several different literatures. Most immediately, our study is related to the literature on information disclosure, as reviewed by Milgrom (2008) and Dranove and Jin (2010). Empirical papers on the effects of information disclosure include Choi, Laibson, and Madrian (2010), Duarte and Hastings (2012), and Duflo and Saez (2003) on financial decisions, Greenstone, Oyer, and Vissing-Jorgensen (2006) on securities, Bhargava and Manoli (2013) on takeup of social programs, Jin and Sorensen (2006), Kling *et al.* (2012), and Scanlon *et al.* (2002) on health insurance plans, Jin and Leslie (2009) on restaurant hygiene, Pope (2009) on hospitals, Bollinger, Leslie, and Sorensen (2011) and Luo *et al.* (2012) on health and nutrition, Dupas (2011) on HIV risk, Figlio and Lucas (2004) and Hastings and Weinstein (2008) on school choice, and many others.

Closely related is the literature that studies the effects of energy-related information. There are a number of papers in this domain that differ from ours in only one respect. Some studies analyze the effects of energy use information disclosure on stated preferences or other proxies for actual purchases of durable goods, including Davis and Metcalf (2014), Deutsch (2010a, 2010b), Newell and Siikamaki (2013), and Ward, Clark, Jensen, Yen, and Russell (2011). Some studies either use observational data (Kallbekken, Saelen, and Hermansen 2013) or randomly assign a very small number of units (Anderson and Claxton 1982). There are other large-sample RCTs that study how peer energy use comparisons affect purchases of durable goods, including Allcott and Rogers (2014), Brandon, List, Metcalfe, and Price (2014), and Herberich, List, and Price (2011), but this social information is conceptually distinct from information about the durable good itself. Also related is Houde's (2014a) analysis of the Energy Star label and other studies of how various kinds of information affect total household energy use, such as Allcott (2011), Dolan and Metcalfe (2013), Jessoe and Rapson (2014), and others. Along with Allcott and Taubinsky (2015), our paper is slightly different in that it uses large sample RCTs to study how providing information about a durable good's energy use affects actual purchases of that good. This particular question is important given the regulatory resources that go into durable good energy use disclosure programs and given the costly energy efficiency standards and subsidies that are partially predicated on the idea that consumers remain imperfectly informed when purchasing durable goods. Furthermore, our paper is substantially conceptually different from the rest of the energy information literature due to its focus on a situation where the information provision process is intermediated by sales agents.

Our experiment is also related to studies of behavior by sales agents and advisers, including field experiments by Anagol, Cole, and Sarkar (2013), Mullainathan, Noeth, and Schoar (2012), and Nagin *et al.* (2002), as well as theoretical analyses by Hoffman, Inderst, and Ottaviani (2013) and Inderst and Ottaviani (2009, 2012). This literature largely focuses on information asymmetries between sales agents and consumers or alternatively on agency problems between firm managers and workers. While these issues could be at play in our context, we do not focus on them. In our model, the firm's inability to observe sales agent behavior does not reduce profits, and agents give consumers correct information about the product. Instead, we highlight agents' imperfect and differential compliance with the experiment. This angle is comparable to findings by Duflo, Gale, Liebman, Orszag, and Saez (2006) that tax preparation professionals have different levels of success in encouraging tax filers to contribute to retirement accounts.

Our experimental tests of sales force incentives connect us to the sales force management literature (see Mantrala *et al.* (2010) for a review and Chan, Li, and Pierce (2014) for a recent example), the "insider econometrics" approach to studying employee compensation and management practices (see Ichniowski and Shaw (2003) for an earlier review), and to "process" field experiments in strategy research (Chatterji *et al.* 2015).

A key innovation relative to the literatures on information disclosure and employee incentives is that we highlight the *interaction* between consumer and firm behavior in equilibrium. Much of the recent empirical work on information disclosure studies consumer responses to information that is experimentally provided with certainty or disclosed by firms under mandate, which isolates consumer behavior independent of the firm. Conversely, much of the employee incentives literature focuses on worker behavior in isolation of the consumer. For example, Ashraf, Bandiera, and Jack (2012), Barankay (2012), Bandiera, Barankay, and Rasul (2005, 2007, 2009, 2010), Larkin (2013), Lazear (2000), Shearer (2004), and others focus on how employees respond to different types of incentives, but these responses do not meaningfully depend on behavior by the firm's customers. By contrast, a central feature of our setting is that the firm's ability to motivate its sales agents to promote a product depends crucially on consumer interest in that product, which is in turn determined by the firm's pricing decisions. A handful of other papers have highlighted other types of equilibrium interactions between the supply and demand sides of the market for energy efficiency, including Fischer (2005, 2011) and Houde (2014b).

3.2 Market Overview

3.2.1 The Water Heater Market

There are over 100 million water heaters currently in use in the United States, approximately one per household.[4] The typical water heater remains in use for 13 years (DOE 2010), which translates into a seven to eight percent replacement rate per year. Non-replacement sales vary with the state of the new housing market, but they historically represent 18 percent of total sales. "Storage" water heaters, where water is stored hot as opposed to heated up on demand,

[4] DOE (2010) provides an overview of the United States water heater market; this is the source of most of the statistics presented here.

make up 96 percent of the market. In 2012, there were 7.69 million residential storage water heaters sold in the United States, 3.96 million (51 percent) of which are fueled by natural gas (AHRI 2013), while the rest use electricity.

Water heater manufacturing is highly concentrated, with the top three producers (A.O. Smith, Rheem, and Bradford White) supplying 96 percent of all residential water heaters in 2008. Approximately half of all units are sold through wholesale distributors; of these, 87 percent are purchased and installed by plumbers. The remaining half of all units are purchased through retail channels such as our partner Retailer. In 2010, the Retailer had a 9 percent share of the retail market, third behind two other retailers that had 23 and 19 percent shares. Thirty percent of the Retailer's sales are made through the call center where our experiment takes place, while the remaining 70 percent are made in physical retail establishments. Our sample thus includes a small but non-trivial share of all water heaters sold in the United States over the study period.

3.2.2 Water Heater Attributes

Water heaters are a convenient product to study because they are differentiated only on a few dimensions. Key characteristics are:

- Fuel type: natural gas, propane, or electric. Consumers' choices depend on what fuels are available in their houses, so choices along this dimension are effectively exogenous. We limit our study to natural gas and propane water heaters, as there are no Energy Star electric models. Less than one percent of sales in the sample were propane-fueled models.

- Storage tank size. This determines the amount of hot water available at any time. Residential tank sizes range from 30 to 80 gallons. In our sample, 90 percent of sales are either 40 or 50 gallons.

- Warranty length. The Retailer offers models with warranty lengths of 3, 6, 9, and 12 years. Models with longer warranties are typically constructed differently, using additional or improved anode rods to delay or fully prevent rusting.

- Tank height. Some consumers need to install water heaters in basements with low ceiling heights. In our sample, about eight percent of sales are "short" models, while the rest are standard height.

- Nitrogen oxide emissions. In some areas, in particular the San Francisco Bay Area and southern California, air quality regulations require consumers to purchase "low NOx" water heaters, which have different natural gas combustion chambers that produce less nitrogen oxide.

- Power venting. Typical natural gas water heaters vent hot combustion gases out of the house through a vertical chimney. When this is not possible, a power vent water heater is required to push the gases outside through a horizontal pipe.

- Mobile home designs. The Retailer sells water heaters specifically designed for mobile homes; these comprise less than one percent of our sample.

- Energy use. We discuss this below.

3.2.3 Water Heater Energy Use and the Energy Star Technology

Each model's energy use is tested at an independent laboratory using test protocols defined by the U.S. government. Test results are used for the FTC Energy Guide labels, or "yellow tags," which are energy use information labels that provide estimated annual energy costs based on national average usage and energy prices. The yellow tags report total energy costs across all fuels used, which for the Energy Star models includes both natural gas and electricity. By law, yellow tags must be displayed on water heaters in a showroom, and the Retailer's website also includes PDFs of the yellow tag next to each model's description. Thus, while many consumers may not see or attend to energy cost information, it is easily verifiable. Of course, each household's actual energy costs may differ from the average due to utilization rates, climate, and other factors.

The energy use test protocols are also used to calculate statistics called Energy Factors, which represent the share of energy input into the water heater that is transformed into hotter water instead of otherwise dissipated. To qualify for Energy Star status, a natural gas water heater must achieve an Energy Factor of 0.67 or above, compared to the standard 0.59.

During our study period, the Retailer sold four natural gas Energy Star models. Two are modified versions of their standard 40 and 50 gallon models with 6-year warranties. To improve energy efficiency, the manufacturer adds another inch of insulation around the tank and uses electric ignition instead of a continuously-burning pilot light. To accommodate electric ignition, the Energy Star models must be plugged in to a power outlet, and they consume a small amount of electricity. They also have electronic thermostats and a more advanced flue damper that opens and closes depending on whether gas is currently being burned. These differences only affect energy use and have no material impact on unit performance, and there are no other differences between the standard and Energy Star 6-year warranty models. (Of course, consumers may attach other connotations to the Energy Star label, and as we mention below, local in-stock availability is an additional differentiator.)

The other two Energy Star models are 40 and 50 gallon premium models with 12-year warranties. The premium models have the same amount of insulation as the standard 12-year models; they achieve higher energy efficiency through electric ignition and other modifications to the combustion process. The premium Energy Star models also differ from the standard 12-year models on other dimensions that make them generally higher-quality.

Table 3.1 presents information on these four Energy Star models and how they compare to their closest substitutes. Standard models cost \$400 to \$700, not including installation. According to the yellow tags, standard models use about \$300 worth of energy each year, meaning that lifetime cost is much larger than upfront purchase price. Energy Star models save about \$30 per year. Because the 6-year warranty models are very close substitutes except for purchase price and energy cost, the 13 and 18 percent internal rates of return are reasonable approximations of the expected monetary net benefits of Energy Star. By contrast, the 12-year Energy Star model also has other premium features, so 1 and 3 percent internal rates of return do not capture the full benefits of the premium model.

Table 3.1: Water Heater Model Overview

	40 Gallon		50 Gallon	
Warranty	6 year	12 year	6 year	12 year
Price ($)				
Standard	420	620	485	665
Energy Star	645	969	700	1020
Annual Energy Cost ($/ year)				
Standard	309	290	315	294
Energy Star	272	261	272	261
Undiscounted payback period (years)	6.1	12.0	5.0	10.8
IRR (at 13 year average life)	13%	1%	18%	3%
Market Share				
Standard	17.6%	6.1%	10.1%	10.4%
Energy Star	0.6%	0.5%	0.2%	0.7%

Notes: This table presents information on the four different Energy Star natural gas water heater models sold by the Retailer, as well as their closest non-Energy Star substitutes. The standard and Energy Star 6-year warranty models are essentially undifferentiated other than price and energy use, while the 12-year warranty Energy Star models have other premium features.

Because of the fixed costs of developing, certifying, and manufacturing a unique model, manufacturers do not produce Energy Star versions of each standard water heater. In particular, the manufacturer that supplies the Retailer does not make mobile home or propane Energy Star models. Because of the fixed costs of procuring and stocking each model, the Retailer also does not carry all the Energy Star models that the manufacturer offers. The Retailer does not carry the manufacturer's short tank size or low-NOx Energy Star models. Because consumers are effectively unable to substitute across these features, we define a subset of "substitutable" models that includes all Energy Star and non-Energy Star natural gas tank water heaters except for low-NOx, short tank height, mobile home, power vent and propane models.

3.2.4 The Sales Process

According to DOE (2010), 35-40 percent of replacement purchases nationwide arise suddenly due to complete unit failure, typically when water rusts through the steel tank and escapes onto the floor. Our follow-up customer surveys show that 83 percent of purchases in our sample were due to unexpected breaks instead of planned replacements. Because most people don't like cold showers, consumers typically want to replace their water heater within 24 hours if possible. This hurry has several implications. First, consumers have not saved money in anticipation of a large expenditure, so they may be especially price sensitive. Second, consumers tend to prefer models that are in stock locally and can thus be installed quickly. Because sales volumes are lower, the Retailer stocks Energy Star water heaters at fewer locations than its standard models. Third, consumers make little time to acquire information about different types of water heaters and their attributes.

A quote from our survey of the Retailer's sales agents nicely summarizes these issues: "Customers that were shopping ahead [i.e. not responding to an unexpected unit failure] seemed to be making more educated decisions ... they were more inclined to use the Energy Star water heaters as item they wanted their quote for. I feel that whenever there was not such a sense of urgency ... customers were in a position to spend more on a better water heater and also able to wait for it to be ordered."

When customers call the Retailer's water heater call center, sales agents have significant influence over their decisions. Some callers have done background internet research and think they know what model they want before calling, while the majority know only that they need a new water heater. Agents work with these callers to determine which model is best for them based on the attributes discussed above, such as fuel type, ceiling height, local low-NOx regulations, and appropriate tank size.

Before the experiment started, we called the Retailer's call center a number of times, acting as "mystery shoppers." We found that the Retailer's sales agents have been successfully trained to look up Energy Star rebates offered by local utilities, discuss Energy Factors, and discuss information on yellow tags. Unless the caller asks about energy efficiency, however, agents never discussed the issue with us, because information disclosure is costly. As one agent wrote on our survey, "I would say about 90 percent of our customers only care about how cheaply can they get away with the purchase of a water heater." Sharing extraneous information increases call times, and many call centers evaluate agents on call times in order to keep labor costs low. Longer call times can reduce customer satisfaction and increase the probability that the customer gets distracted and does not complete the sale.

These features of the water heater sales process motivated our experiment: perhaps Energy

Star sales are low because consumers are unaware of the product and its benefits, and agents' influence over consumers could be leveraged to increase awareness. Notwithstanding, this discussion also makes clear that informing consumers is costly. In the next section, we present a model that captures these issues.

3.3 Experimental Design and Data

3.3.1 Sales Associates and the Sales Process

The Retailer's water heater division operates two call centers. There are 77 sales agents who take at least one call during our sample period. These sales agents sell only water heaters, not other goods. The agents report to Team Managers, who in turn report to Shift Managers, who report to the call center manager. Agents make between \$11 and \$14 per hour, depending on seniority, along with sales incentives that typically scale closely with purchase price and average approximately \$4. Interestingly, however, sales incentives are only slightly higher for the 6-year warranty Energy Star models compared to their closest non-Energy Star substitutes, despite purchase prices that are about \$200 higher. Sales incentives for the premium 12-year warranty Energy Star models are about \$5 larger than for the standard 12-year warranty non-Energy Star models.

The Retailer has an established set of processes that sales agents are to follow on each call. About 60 to 65 percent of calls are recorded at random, and managers monitor a subset of these calls for evaluation and quality assurance. The sales agents meet with their managers weekly to review performance and talk about sales initiatives and modifications to the sales process.

When a customer calls, he or she is routed to the first available sales agent. The call centers use caller ID, and the agents verbally confirm the caller's phone number. Using this phone number, the customer is assigned a "reference number." We define a "consumer" as a unique reference number. Individuals often call more than once as they comparison shop or gather more information. If an individual calls more than once from the same phone number or verbally gives the same number to a sales agent, then he or she is tracked as a unique customer.

Once the sales agent and consumer agree on a water heater model, the sales agent checks whether the model is in stock in the customer's region, arrives at a price quote, records the customer address, and charges the customer's credit card. Customers can install the unit themselves, hire a third-party plumbing contractor, or pay the Retailer to do the installation.

3.3.2 Experimental Design

The sales agents have a standard computer interface that has the Retailer's sales program plus internet access. To implement the experiment, the Retailer's staff redesigned the interface to open the experiment website each time a reference number is entered. Our research team designed and programmed the experiment website, which afforded us full control over the randomization and other content.

Appendix Figure A.1 is a screen shot of the website's initial screen. On this screen, the agent would enter the customer's needed fuel type (Gas or Electric) and click "GO." After the agent clicked "GO," the website would display call handling instructions, including a script that the agent was to cover with the customer. The different treatments were implemented through these scripts. Appendix Figure A.2 is a screen shot of one example treatment, the $100 Energy Star rebate. Electric customers are excluded from the experimental population, with the website displaying "No Script." Natural Gas customers form the experimental population, and they are randomly assigned to one of the treatment groups.

Agents and callers were both randomized. Agents were randomly assigned as Information Treatment Agents or Information Control Agents. Callers were randomly assigned to treatment groups based on their reference number. Thus, consumer who called multiple times but kept the same reference number remained in the same treatment condition. Consumers who first spoke with an Information Control Agent were automatically assigned to Information Control, while callers who first spoke with an Information Treatment Agent were randomly assigned to either Information Treatment or Information Control.

Table 3.2 displays the experimental timeline and treatment groups. Phase 1 ran from November 2012 to April 2013. During this time, there was a three-by-two matrix of treatments: customers were randomly assigned to $0, $25, or $100 rebate, which was crossed with Information Treatment or Control. Phase 2 of the experiment ran from early April to early June 2013. In this phase, we added a sales incentive, which the Retailer calls a "spiff." Phase 3 lasted from early June to early July. In this phase, we added two final treatments, which were interactions of the spiff with the two rebate levels. In Phase 4, the Retailer ended the spiff treatments but continued the rest of the experiment for several weeks. In total, there were eight different treatment cells, plus control.

Table 3.2: Experiment Timeline

Phase	Dates	Info, Rebates, and Info x Rebates	Spiff	Spiff x Rebates	Consumers in Sample	Sales
1	Nov 21-April 6	Yes			12,629	4,675
2	April 7-June 13	Yes	Yes		7,254	2,523
3	June 14-July 6	Yes	Yes	Yes	1,974	715
4	July 7-July 26	Yes			1,490	362

Below, we give examples of the call handling instructions for several example treatments. In the Information Treatment condition with no rebate, the website instructed the agent to read the following script to the customer:

> *Let me take a moment to tell you about our Energy Star models. Energy Star water heaters cost about $220 more than a standard model, but they save a typical household $40 each year, so you would make up that price difference in about six years. Over 12 years, which is the normal life of a water heater, you would save $480 in energy bills. Energy Star models may not be available for every home. If possible, would an Energy Star water heater be of interest to you?*

In the rebate condition with no information, the agent was instructed to say:

> *I have good news. [Retailer] has specially selected you for a $100 rebate on any Energy Star water heater. Energy Star models may not be available for every home. If possible, would an Energy Star water heater be of interest to you?*

If the customer was assigned to the spiff, the call handling instructions read:

> *ENERGY STAR SPIFF CALL*
>
> *You (the Retail Hotline Associate) will receive $25 on your next paycheck if this caller buys any Energy Star water heater. You can share with the caller any useful information about the benefits of Energy Star, perhaps including environmental benefits or long-run energy cost savings. The caller does not need to purchase on the*

initial call. If the same caller calls back later and uses the same reference number, all RHAs that spoke with that reference number earn the $25.

In the combined spiff plus rebate conditions which were added in Phase 3, the sales agent was not instructed to read a specific script. Instead, the call handling instructions told sales agents that the customer was eligible for a rebate and left it to the sales agent to decide how to phrase that information.

> *ENERGY STAR SPIFF CALL + $25 CUSTOMER REBATE*
>
> *You (the Retail Hotline Associate) will receive $25 on your next paycheck if this caller buys any Energy Star water heater.*
>
> *The customer will also receive a $25 rebate off of any qualifying Energy Star model.*

In the control group, the instructions read:

> *CONTROL GROUP: NO SCRIPT*
>
> *This customer is in the control group. Proceed with the call as you normally would. Answer any questions the customer has, but try not to use any of the language in the information treatment script.*

At the end of the call, the sales agents reported in the experiment website whether or not they delivered the script. As we shall see, these self-reports overstate compliance relative to our independent audits. The website and the team managers instructed the agents that the only reasons not to deliver the script were if the customer needed a low-NOx, short tank, or other specialty model that was not substitutable with Energy Star models. If the agent did not complete the script, the website required them to select the reason for non-compliance from a dropdown menu.

The experiment was closely integrated into the call center processes. At the outset, managers trained the agents on the scripts and how to use the website, and this was also part of training for newly-hired agents during the experiment. We also communicated directly with the agents through several group emails and two videos that explained the importance of compliance with the experiment. Specifically, we emphasized the importance of both delivering the scripts on treatment group calls and *not* discussing elements of the scripts on control group calls.

Every week of the experiment, we provided the Retailer with agent-specific compliance reports based on self-reported compliance from the website. We had bi-weekly calls with managers to discuss these compliance reports, and managers could then discuss with agents in their weekly meetings. The Retailer's internal call monitors also audited calls for compliance with treatment assignment. Agents with low compliance with the experiment were pressured by managers to do better. To encourage competition between the two call centers, managers also reviewed average compliance for each call center, as well as trends over time. In the endline survey, sales agents reported that managers frequently emailed and talked with them about the experiment. In sum, agents did face some costs if they did not at least report compliance with the experimental protocols. However, this experiment was only one of many issues that managers and agents needed to attend to.

Individuals who call multiple times from multiple phones and do not tell the sales agents that they have previously called would have been assigned different reference numbers, and thus potentially different treatments. This could generate spillovers, for example if a caller who purchases using an Information Control reference number had been assigned an Information Treatment reference number on a previous call. Based on our conversations with Retailer staff, we do not believe that this happens on more than a handful of calls, although we do not have a precise estimate.

Some consumers, perhaps plumbing contractors or landlords of multiple homes, order multiple water heaters during the experiment. The Retailer gives these individuals separate reference numbers on their separate purchases, and they are thus treated as separate "consumers" in the experiment. While this also could generate spillovers, it could not have more than a negligible impact on the estimates because it affects only a very small share of the sample: there are 104 individuals, or 0.4 percent of the sample, who order two water heaters from the same phone number, and no phone number appears more than twice in the sales data.[5]

[5] Other than perhaps these 104 individuals, the consumers that call the call center are the final owners of the water heater; contractors do not order through the retail hotline. Furthermore, consumers typically are not already working with a contractor before calling the Retailer, because a contractor would typically procure the water heater on behalf of a consumer.

3.3.3 Data

There are several main data sources. The first is the Retailer's call database. An observation consists of the unique customer reference number, date and time of the call, and the agent receiving the call. This database includes only sales calls, not warranty service, repairs, or other types of calls.[6] Using the reference number, this is matched to the Retailer's purchase data, which include the model purchased, price paid, and other details.

The Retailer's call database is also matched by reference number to the experiment website database. This database includes the treatment assignment and the agent's self-reported compliance, for each reference number where the website was opened. For the 1.3 percent of reference numbers that appear in the website data multiple times, we code that the script was read to the consumer if any agent reported that he or she had done so on any call. In the regressions, each consumer i must be associated with an individual agent a; we use the last agent in the website who spoke with customer i.[7] We define a variable N^s_{iat} that takes value 1 if agent a reported compliance on a treatment group call with consumer i, and 0 otherwise. Agents were not explicitly asked to read a script on spiff treatment calls or on controls calls. We define N^s_{iat} as missing for spiff treatment calls and zero for control calls.

The total number of consumers (reference numbers for consumers interested in natural gas water heaters) recorded in the call database during the experiment is 38,179. Of these, 23,347 (61 percent) are in the website and are randomly assigned to a treatment group; these calls comprise our "experimental population." The calls that are not recorded in the website are largely conversations that did not last long enough for the sales agent to activate the website. As Table 3.2 shows, 35 percent of consumers (8275 in total) purchased from the Retailer. Of these sales, 73 percent were either Energy Star models or were substitutable with Energy Star models given other attributes such as tank height and NOx emissions. Of this substitutable group, however, only 3.5 percent were actually Energy Star, while the remaining 96.5 percent were standard models.

[6] Approximately 2-3 percent of reference numbers are repeated, typically as the sales agent updates information. In these cases, we use the most recent observation. We drop six reference numbers that appear to be used twice for two distinct individuals.

[7] There are other ways to code this, but it would not matter, because almost all of the 1.3 percent of reference numbers that appear multiple times were entered by the same agent. Only four reference numbers were entered by two separate agents in the website, and none are entered by 3 or more. Because treatment groups were assigned by reference number, a consumer's treatment assignment is maintained even if an agent enters the same reference number multiple times.

For consumers that purchased water heaters, the Retailer records their name and address. Zip codes were used to match median income from the most recent American Community Survey (ACS) 5-year estimates and the the share of vehicles registered in the zip code that are hybrids, which could be an important correlate of environmentalism and interest in energy efficiency. Using each purchaser's name and address, a marketing data company called Acxiom provided assessed home value, college graduate indicator, age, household size, and political affiliation. Acxiom gathers data from public records, magazine subscriptions, voting records, scanner data, online purchases, and other sources, and their data are certainly measured with error. For the approximately 10 percent of addresses missing the college graduate indicator, age, or household size, we substitute zip code-level means, again from the ACS 5-year estimates.

Using the Acxiom political affiliation data, we construct a variable called "Democrat" which takes value 1 if the purchaser is a registered democrat, 0 if a republican, and 0.5 if independent or unregistered. If political affiliation is missing, we replace Democrat with the county-level ratio of Democrat to Democrat plus Republican votes in the 2004 and 2008 presidential elections, using data from the U.S. Election Atlas (Leip 2013). Acxiom also provided two additional levels of environmentalism. "Environmentalist" is an indicator variable for whether the consumer subscribes to environmental magazines or contributes to environmental or animal welfare charities. "Green Living" is an indicator that takes value one if Environmentalist equals one or if the household purchases environmentally-healthy products such as eco-friendly soaps and organic foods.

Table 3.3 presents sample means and standard deviations for our nine demographic variables. Consumers in our sample are older and wealthier than the general population, likely related to the fact that they are almost entirely homeowners. They are also more liberal and environmentalist, as illustrated by their Democrat scores and zip code hybrid vehicle shares. In our data, Energy Star demand is positively associated with zip code median income, home value, and zip code hybrid share, reminiscent of the findings of Kahn (2007). This suggests that the Energy Star market share would be even lower in a nationally-representative sample.

Table 3.3: Representativeness

	Sample Mean	Sample Std. Dev.	National Average
Zip Median Income (000s)	71	27.3	56.9
Home Value (000s)	338	293	246
College Grad	0.61	0.43	0.32
Age	57.3	13.4	37.3
Household Size	3.2	1.5	2.4
Democrat	0.62	0.34	0.53
Zip Hybrid Share (out of 100)	1.3	1.1	0.94
Acxiom Green Living	0.31	0.46	-
Acxiom Environmentalist	0.14	0.35	-

This table gives the mean and standard deviation of customer demographics. These variables are matched based on addresses and are thus available only for consumers who purchase water heaters. National average college graduate share is for people older than 25 years. National averages for the Acxiom Green Living and Environmentalist variables are not available.

We do not have much information with which to test for balance on observables within the entire experimental population. Using the address-based demographic information, however, we can test for balance within the subset of consumers who purchased from the retailer. Appendix Table A.1 presents tests using the nine demographic variables for each of the eight treatment cells, relative to control. Only one of the 72 t-tests rejects equality with greater than 90 percent confidence, and all nine F-tests fail to reject that the treatment groups are balanced on observables.

3.3.3.1 Customer Follow-Up Surveys

We hired an independent call center to conduct telephone follow-up surveys of customers who had called between December 10 and June 29. We designed two separate surveys, one for consumers who had purchased from the Retailer and one for consumers who had called but not purchased. For purchasers, we asked a battery of questions covering household information, the water heater purchase process, and the Energy Star product. For non-purchasers, we asked whether they had purchased an Energy Star water heater and why they had decided not to buy from the Retailer. For this analysis, we focus on questions related to consumers' knowledge of the Energy Star model, which we only asked of purchasers. Any other results are certainly available upon request, and the survey protocols are available as part of the replication files.

We directed the call center to complete no more than 200 surveys of purchasers and as many surveys as possible of purchasers. In order to maximize response rates, we offered a $25 gift card from the Retailer to any respondents who initially attempted to refuse; 149 people accepted. In total, there were 1,091 completed surveys (including 891 from purchasers) from 6,342 attempts, for a response rate of 17 percent.

3.3.3.2 Independent Audits of Recorded Phone Calls

Our research assistant (RA) audited 2,122 calls from natural gas water heater consumers recorded between May 1 and July 18. These 2,122 calls are all recorded calls assigned to any of the treatment groups during that period, along with approximately five calls per day from the control group. The audits were blind, meaning that the RA did not know the treatment assignment when auditing a call.

There are two reasons why a call is not observed in our audit data. First, the Retailer's software records only a randomly-selected 60 to 65 percent of calls. Second, the database of recorded calls is not organized by reference number, so our RA needed to match recordings to reference numbers using phone number, time and duration of call, and other information; not all calls could be matched.

We worked with the RA to develop a protocol for quantifying the content of the interaction between the sales agent and the customer.[8] We measure the information provision process using six variables:

- *Mentioned E-Star*: Did the agent mention energy efficiency, energy use, or Energy Star?
- *Rebate*: Did the agent mention the experiment's Energy Star rebate from [the Retailer]?
- *Saves Money*: Did the agent mention that an Energy Star (or energy efficient) water heater saves money in energy costs?
- *Payback Period*: Did the agent quote a payback period?
- *Read Script*: Does the agent say one of the experiment scripts, either exactly or approximately?

[8] By necessity, this was done before writing the paper, as the Retailer deletes call recordings after 30 days.

- *ln(E-Star Seconds)*: For how many seconds did the agent and the customer talk about energy efficiency, energy use, or Energy Star? We use the natural log of one plus this number.

Because there are a small share of consumers who call multiple times, the audit dataset includes multiple observations of some consumers. Thus, there are 2,069 unique consumers in our experimental population for whom we have audit data. For the five binary variables above, we collapse using the maximum. In other words, consistent with our construction of agents' self-reported compliance N^s_{iat}, we measure whether a consumer was ever informed about Energy Star or a rebate. For the sixth variable, the number of seconds discussing Energy Star, we take the sum across all of a consumer's calls.

3.4 Model

We present a simple model of information disclosure through agents in an imperfectly competitive market. The model is related to other models of informative and persuasive advertising, such as Becker and Murphy (1993), Dixit and Norman (1978), Dorfman and Steiner (1954), Kotowitz and Mathewson (1979), and Shapiro (1980). It builds most closely on the simple spatial competition model of Hotelling (1929) and the Grossman and Shapiro (1984) extension to informative advertising. We extend these models by modeling two goods, a base good and a high-quality good, and by nesting sales agents within the model, which provides microfoundations for the experiment's information disclosure process.

The model by itself would not be an important contribution, and what we call "propositions" are straightforward comparative statics. We include the model, however, because it helps to put the empirical work in context. The sales agent model generates predictions for the field experiment, under the realistic assumption that competing firms do not respond to our randomly-assigned treatments. We then solve for the two-firm equilibrium, which highlights potential policy implications.

3.4.1 Setup

3.4.1.1 The Choice Set

Two firms are located at either end of a Hotelling line. Aside from their locations, the firms are fully symmetric. Each firm sells two goods indexed by j: a base good I and a high-quality good E. For applications to energy-using durables such as water heaters, we have in mind that good E is energy efficient and good I is energy inefficient.[9] The two goods are produced at marginal costs c_j, and firm f sells good j at price p_{fj}. Consumers have unit demand, meaning that they will buy one of the two goods from one of the two firms. The unit demand assumption is certainly appropriate for water heaters. The basic Hotelling setup also fits the water heater market very well, with horizontally-differentiated retailers selling closely-comparable sets of products.

3.4.1.2 Consumers

A group of consumers normalized to measure 1 is distributed uniformly along the Hotelling line, with locations indexed by $d \in [0, 1]$. The variable t measures "transport costs" - the utility of purchasing from the closer vs. more distant firm. Of course, d could reflect brand loyalty or many factors other than geographic distance. There are two consumer types, High and Low, indexed H and L, in proportions λ and $1-\lambda$, respectively. Both types earn gross utility V from owning the base good. A consumer at point d realizes net utility $V - dt - p_{1I}$ if purchasing good I from firm 1 and $V - (1-d)t - p_{2I}$ if purchasing from firm 2.

When owning good E, low types earn gross utility V_L and high types earn utility V_H, with $V_H > V_L$. In applications to energy-using durables, we have in mind that type H has high utilization demand or other reason to prefer energy efficiency. A High type consumer at point d realizes net utility $V_H - dt - p_{1E}$ if purchasing good E from firm 1 and $V_H - (1-d)t - p_{2E}$ if purchasing from firm 2. Net utility for Low type consumers purchasing good E is analogous, except with V_L instead of V_H.

Define $G_{HE} = V_H - c_E$ and $G_{LE} = V_L - c_E$ as the surplus from allocating good E to type H and type L, respectively, and define $G_I = V - c_I$ as the surplus from allocating good I to

[9] An energy efficient good may or may not be "high-quality" on other dimensions, but it is vertically differentiated in the literal sense that all consumers should prefer lower energy costs.

either type. We assume that $G_{HE} > G_I$ and $G_{LE} < G_I$, meaning that it is socially optimal to for High type consumers to purchase good E and for Low type consumers to purchase good I.

Consumers are initially unaware that the high-quality good exists, and they become aware only if informed by a sales agent. Sales agents at firm f inform consumers only about good E sold by firm f. For example, if a consumer is informed by an agent of firm 1, then he or she is aware only that firm 1 sells good E. Consumers informed by both firms are aware that both firms sell good E. There are thus four different types of consumers. Shares $\theta_1(1-\theta_2)$ and $\theta_2(1-\theta_1)$ are informed only by firm 1 or firm 2, respectively. Share $\theta_1\theta_2$ is informed by both firms, while share $(1-\theta_1)(1-\theta_2)$ is informed by neither. We focus on the set of parameter values under which equilibrium prices are such that Low type consumers always purchase good I and High type consumers purchase good E if and only if they are informed.

Define q_{fj} as the quantity of consumers that purchase good j from firm f. Denote $d_E = d_E(p_{1E}, p_{2E}, t)$ and $d_I = d_I(p_{1I}, p_{2I}, t)$ as the largest d such that the consumer at that location purchases good E and I, respectively, from firm 1. Consumers to the "left" of d_E and d_I (on a horizontal Hotelling line) purchase from firm 1, while consumers to the "right" purchase from firm 2.

Low types and uninformed High types purchase good I, and they represent share $(1 - \lambda(\theta_1 + \theta_2 - \theta_1\theta_2))$ of the population. Firm 1's quantity sold of good I is thus $q_{1I}(p_{1I}, \theta_1) = d_I(1 - \lambda(\theta_1 + \theta_2 - \theta_1\theta_2))$. Informed High types purchase good E. Share $\theta_1(1-\theta_2)$ have only been informed by firm 1, and they thus purchase good E from firm 1 with certainty. Share $\theta_1\theta_2$ have been informed by both firms, and such consumers purchase from firm 1 if they are located to the left of d_E. Thus, firm 1's quantity sold of good E is $q_{1E}(p_{1E}, \theta_1) = \theta_1(1-\theta_2) + \theta_1\theta_2 d_E$. Quantity equations for firm 2 are symmetric, except replacing d_I and d_E with $1-d_I$ and $1-d_E$, respectively.

3.4.1.3 Sales Agents and Information Provision

The firm employs a number of sales agents normalized to measure 1. As described above, providing information takes time and effort, and it reduces sales agents' ability to successfully carry out other required tasks during sales interactions. We assume that informing share θ_f of consumers costs firm f's sales agents $\frac{\alpha}{2}\theta_f^2$, where α is a scaling parameter. In the context of our field experiment, such costs to sales agents are necessary to explain why they sometimes do not deliver information when the experiment protocol directs them to. Convex costs are needed to generate a unique equilibrium, but this convexity also informally captures another important feature of our empirical results: as we shall see, sales agents are able to target their attempts at information provision at consumers who appear to be more receptive. This suggests that as sales agents attempt to inform more consumers, the marginal consumers are increasingly unreceptive.

The firm does not observe all individual sales interactions, so the firm cannot directly set θ_f. Instead, firms pay sales agents a fixed wage w_f plus sales incentive s_f for each sale of good E. Sales agents are risk-neutral, and they maximize the following utility function:

$$U(\theta_f) = s_f q_{fE}(p_{fE}, \theta_f) - \frac{\alpha}{2}\theta_f^2 + w_f \tag{3.1}$$

Agents have an outside option which gives utility $\overline{U}$, which gives the participation constraint that $U(\theta_f) \geq \overline{U}$.

3.4.2 Sales Agent Information Provision Behavior

Agents inform consumers until the marginal incentive pay equals the marginal cost of information provision. Plugging in for q_{1E} and maximizing agents' utility gives agents' utility-maximizing choice of information provision, which we denote as $\tilde{\theta}_f$. For firm 1, this is:

$$\tilde{\theta}_1 = \frac{s_1}{\alpha} \cdot \frac{\partial q_{1E}(p_{1E}, \tilde{\theta}_1)}{\partial \theta_1} = \frac{s_1}{\alpha}\lambda \left[(1-\theta_2) + \frac{p_{2E} + t - p_{1E}}{2t}\theta_2\right] \tag{3.2}$$

The equation for firm 2 is symmetric.

In our empirical work, we study how changes in price and sales incentives affect agents' information provision and consumers' purchases. Because the experimental treatment groups are a small share of the overall market, the results are likely to be informative about the best response functions for the Retailer and its sales agents, not equilibrium effects after competing firms re-optimize. The model generates three very simple comparative statics, which we label as Propositions 1-3.

Proposition 1 covers the initial sales agent behavior, independent of any experimental manipulations. It states that agents inform consumers more when potential demand from informed consumers is larger.

Proposition 1 $\frac{\partial \tilde{\theta}_f}{\partial p_{fE}} < 0$ *and* $\frac{\partial \tilde{\theta}_f}{\partial \lambda} > 0$.

Proposition 1 follows immediately from Equation (3.2). Intuitively, the proposition states that agents do not bother to inform consumers if information doesn't increase demand. This could occur either if the price is high or if most consumers are Low types.

Proposition 2 covers agents' response to changes in the sales incentive.

Proposition 2 $\frac{\partial \tilde{\theta}_f}{\partial s_f} > 0$. *Furthermore,* $\frac{\partial^2 \tilde{\theta}_f}{\partial s_f \partial p_{fE}} < 0$ *and* $\frac{\partial^2 \tilde{\theta}_f}{\partial s_f \partial \lambda} > 0$.

The first part of Proposition 2 states the straightforward result that agents provide more information about good E if incentivized to sell good E. The second part states that if information doesn't increase demand very much, sales incentives do not induce agents to provide much more information. Put simply, agents are unresponsive to sales incentives if consumers don't respond to agents' actions. Both parts of this proposition follow from taking derivatives of Equation (3.2).

Our experiment varies both high-quality good price p_{fE} and sales incentive s_f. Should these have independent effects on quantity demanded, or should a lower price reinforce the impacts of a larger sales incentive?

Proposition 3 $\frac{\partial^2 q_{fE}}{\partial p_{fE} \partial s_f} < 0$.

Proposition 3 states that the effects of lower prices and higher sales incentives reinforce each other. When prices are lower, the potential quantity demanded from High types is higher, making sales agents more responsive to increases in the sales incentive. When the sales incentive is higher, sales agent behavior is more responsive to changes in potential quantity demanded that result from lower prices. This follows from taking derivatives of $q_{fE}(p_{fE}, \tilde{\theta}_f)$; see Appendix B.1.

3.4.3 Market Equilibrium

3.4.3.1 The Profit Function

Firms set prices, sales incentives, and wages to maximize profits. The profit function for firm 1 is:

$$\pi_1(p_{1I}, p_{1E}, s_1, w_1) = q_{1I}(p_{1I}, \tilde{\theta}_1)\,[p_{1I} - c_I] + q_{1E}(p_{1E}, \tilde{\theta}_1)\,[p_{1E} - c_E - s_1] - w_1 \tag{3.3}$$

Substituting the sales agents' participation constraint into the profit function (and dropping the tilde on θ for simplicity) gives:

$$\pi_1(p_{1I}, p_{1E}, \theta_1) = q_{1I}(p_{1I}, \theta_1)\,[p_{1I} - c_I] + q_{1E}(p_{1E}, \theta_1)\,[p_{1E} - c_E] - \frac{\alpha}{2}\theta_1^2 - \overline{U} \tag{3.4}$$

The fact that Equation (3.3) reduces to Equation (3.4) implies that when setting the sales incentive, the firm acts as if it maximizes profits net of agents' information costs. Thus, although the firm does not explicitly contract on agents' information provision θ_f, this does not generate inefficiency. Because both firm and agent are risk neutral and the agent does not have limited liability, the firm does not need to insure the agent, and the firm's choice of s_f will maximize the sum of profits and agent utility. Thus, there is no "agency problem" of the traditional sort.

3.4.3.2 Equilibrium

Appendix B.1 solves for the symmetric Nash equilibrium. As in the standard textbook model, equilibrium prices are $p^*_{fI} = c_I + t$ and $p^*_{fE} = c_E + \left(\frac{2}{\theta^*_f} - 1\right) t$. Notice that unless $\theta^*_f = 1$, $p^*_{fE} - c_E > p^*_{fI} - c_I$, as imperfect information about competitors' product availability gives firms additional market power when selling good E.

The equilibrium information provision level is $\theta^*_f = \frac{-5+\sqrt{32\frac{\alpha}{\lambda t}-7}}{4\left(\frac{\alpha}{\lambda t}-1\right)}$. This is decreasing in information cost α. It is increasing in the share of High types λ, as more High types increase the chance that an informed consumer will purchase good E, which has higher markup. Intuitively, if there are not many High types in the population, firms will not try hard to inform all consumers just to change decisions by the small number of consumers who might purchase good

E. Information provision is also increasing in t, because higher transport cost increases the relative markup on good E, which increases the returns to informing consumers about that good.

3.4.4 Social Welfare

Social welfare can be written as surplus from High types plus surplus from Low types net of information provision costs and transport costs. In the symmetric equilibrium where Low type consumers always purchase good I and High type consumers may purchase either good depending on θ_f, social welfare can be written as:

$$SW(\theta_f) = \lambda \left[(2\theta_f - \theta_f^2)G_{HE} + (1-\theta_f)^2 G_I\right] + (1-\lambda)G_I - \alpha\theta_f^2 - \frac{t}{4} \tag{3.5}$$

Because consumers have unit demand, there is no extensive margin distortion from oligopoly pricing: although the market is not perfectly competitive, prices do not distort total quantity demanded. Furthermore, because firms are symmetric, they offer the same price in equilibrium, and there is no way to reduce the losses from transport costs. Therefore, distortions arise only when consumers purchase a good which is not optimal given their type - in particular, when High types do not purchase good E because they are uninformed.

Appendix B.1 derives the socially-optimal level of information provision θ^+ and shows that it could be greater than, less than, or equal to the market equilibrium level. The market is more likely to under-provide information when the social gain $G_{HE} - G_I$ from moving consumers to good E is large or when market power t is small.

We assume that the government cannot directly mandate sales agents or firms to achieve θ^+, which seems realistic in most contexts. Instead, consider a government that has two potential policy instruments: a minimum quality standard or a sales incentive for good E provided directly to the agent.[10] What are the welfare implications?

[10] Our model is not well-suited to consider a subsidy for good E. In our model, a subsidy for good E is equivalent to an equal reduction in marginal cost for both firms. As long as the subsidized price still exceeds p^*_{fI}, firms pass the subsidy on to consumers in the form of reduced prices, but because the price reduction is symmetric, there is no change in θ^*. Thus, no additional consumers are informed, and the market equilibrium is unaffected except for the transfer from taxpayers to the government to firms to some High types. This result would be different if we had a continuum of quality preferences instead of just binary H and L types. Note that in Propositions 1 and 3 and in our experiment, only one firm reduces prices, so $\tilde{\theta}_f$ and q_{fE} do respond to changes in p_{fE}.

3.4.4.1 Welfare Effects of a Minimum Quality Standard

As discussed in the introduction, minimum energy efficiency standards are often justified by the assertion that imperfect information reduces demand for energy efficient goods. In this model, one can capture the effects of a minimum quality standard by eliminating the base good and requiring that all consumers purchase good E. Under our assumptions about G_{HE}, G_{LE}, and G_I, this increases welfare for High type consumers but decreases welfare for Low types. A minimum quality standard also eliminates information provision costs, as all consumers now are mandated to purchase good E. Under a minimum quality standard, social welfare is:

$$SW^{standard} = \lambda G_{HE} + (1-\lambda)G_{LE} - \frac{t}{4} \tag{3.6}$$

Subtracting the $SW(\theta_f^*)$ as written in Equation (3.5), we have the welfare gains of a standard relative to the market equilibrium:

$$SW^{standard} - SW(\theta_f^*) = \lambda(1-\theta_f^*)^2(G_{HE} - G_I) + (1-\lambda)(G_{LE} - G_I) + \alpha\theta_f^{*2} \tag{3.7}$$

The first term is gains to the share of High types that are uninformed in the market equilibrium. The second term is negative, capturing losses from Low types who must now purchase good E. The third term is savings in information provision costs. Welfare gains are mechanically larger when the social gains from good E are larger and when there are more High types. Welfare gains are also larger when α is large, i.e. when it is costly to inform consumers.

3.4.4.2 Welfare Effects of Government-Provided Sales Incentives

Unlike minimum quality standards, it would be novel for governments to provide sales incentives directly to sales agents. What effects would this have in this model? Consider the case where firms under-provide information relative to the social optimum (i.e. $\theta^+ > \theta^*$), so the government would like to subsidize information provision. Assume that the government can fund the subsidy with lump-sum taxes.

If firms can set $s < 0$ (or provide a higher sales incentive for good I), then government-provided sales incentives act just like a marginal cost reduction, and firms will offset the government incentive to arrive at their profit maximizing level of information provision in Equation (B.6). However, consider the potentially-realistic case where the firm cannot set $s < 0$, i.e. it cannot pay its agents to *not* sell a product or otherwise undo the government-provided sales incentive. In this case, the government will set the sales incentive that induces the socially-optimal information provision effort, and the firm will provide no additional sales incentive and will adjust the fixed wage such that agents' participation constraint binds. Substituting θ^+ into the agents' first-order condition in Equation (3.2) and simplifying, this gives the socially-optimal government-provided sales incentive:

$$s^+ = \frac{\alpha\theta^+}{\lambda(1 - \frac{1}{2}\theta^+)} \tag{3.8}$$

Increased information provision also has the side effect of increasing competition in the market for good E, so both firms decrease their prices p^*_{fE}. In this model, this has no welfare implications, but it does generate a transfer from firms to consumers.

3.4.5 Discussion

The model helps to understand and motivate the field experiment that we detail in the remainder of the paper. The three simple propositions show that sales agent behavior is not only important in directly determining the firm's ability to disclose information, but also indirectly informative about consumers' preferences: because information provision is costly, optimizing agents will not inform consumers if they believe that information will not affect demand. The social welfare discussion highlights that if sellers' information disclosure is not effective at increasing demand for high quality goods but the social benefit of allocating good E to High types $(G_{HE} - G_I)$ is large, minimum quality (energy efficiency) standards are more likely to increase welfare. Before drawing this conclusion, however, we would need to show that information provision is ineffective because disclosure is difficult (i.e. α is large), not because consumers still don't want good E after being informed (i.e. λ is small). The model also highlights a novel policy that could increase welfare: government-provided sales incentives. To evaluate this potential policy, one similarly needs to know α and λ, along with $(G_{HE} - G_I)$.

3.5 Experiment Results

3.5.1 Sales Agent Behavior

Define N_{iat} as a measure of whether agent a provides information to consumer i during phase t of the experiment. We observe sales agent behavior from two sources: self-reported compliance N^s_{iat} and the independent audits. While the latter measure is only available for a smaller subset of calls, it is an independent assessment and also provides multiple measures of what the sales agent said. $\mathbf{T}_i$ is a vector of indicator variables for each of the eight treatment cells. We estimate how N_{iat} varies across treatments using the following equation, where ϕ_t is a vector of indicators for the four phases of the experiment, μ_a is an agent fixed effect, and υ_{iat} is the error term:

$$N_{iat} = \beta \mathbf{T}_i + \phi_t + \mu_a + \upsilon_{iat} \tag{3.9}$$

The ϕ_t and μ_a controls are necessary because assignment probabilities vary across phases as we added treatment groups and across agents who were assigned to Information Treatment or Control. This equation is estimated as a linear probability model (LPM) in OLS with robust standard errors. In typical cases like ours where the true probability model is not known, Angrist and Pischke (2012) advocate for using the LPM instead of an arbitrary non-linear model such as probit or logit, and we follow their recommendation. In practice, our results are qualitatively and quantitatively very similar when using probit, logit, or the LPM.

Column 1 of Table 3.4 presents the results using self-reported compliance N^s_{iat} as the measure of compliance. The sample excludes the spiff treatment calls because agents were not explicitly asked to read a script on these calls. Agents report that they read the script on 46 to 49 percent of calls, and this depends little on treatment assignment. Columns 2-7 use data from our independent audits, showing that self-reports substantially overstate compliance. Relative to control group calls, agents were about ten percent more likely to mention Energy Star in Information Only treatment group calls, and about 14 percent more likely to do so in Rebate Only calls.

Columns 3-6 directly measure compliance with the experiment scripts. Column 3 shows that agents mentioned the experiment's Energy Star rebate on 14 to 24 percent of calls when

Table 3.4: Compliance by Treatment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable:	Agent Reported	Mentioned E-Star	Rebate	Saves Money	Payback Period	Read Script	ln(E-Star Seconds)
1(Information Only)	0.481	0.108	0.028	0.117	0.133	0.149	0.406
	(0.010)***	(0.037)***	(0.013)**	(0.031)***	(0.024)***	(0.027)***	(0.133)***
1(Info and 25 Rebate)	0.491	0.130	0.239	0.168	0.149	0.193	0.516
	(0.010)***	(0.039)***	(0.030)***	(0.033)***	(0.025)***	(0.033)***	(0.141)***
1(Info and 100 Rebate)	0.460	0.057	0.143	0.108	0.093	0.156	0.337
	(0.019)***	(0.060)	(0.045)***	(0.052)**	(0.036)***	(0.057)***	(0.232)
1(25 Rebate Only)	0.491	0.147	0.182	0.019	0.000	0.214	0.509
	(0.006)***	(0.025)***	(0.018)***	(0.016)	(0.005)	(0.020)***	(0.089)***
1(100 Rebate Only)	0.494	0.143	0.190	0.061	0.001	0.180	0.620
	(0.010)***	(0.037)***	(0.031)***	(0.027)**	(0.009)	(0.033)***	(0.144)***
1(Spiff Only)		0.027	0.032	0.008	-0.002	0.011	0.099
		(0.024)	(0.010)***	(0.016)	(0.005)	(0.013)	(0.085)
1(Spiff and 25 Rebate)		0.051	0.092	0.073	-0.014	0.036	0.257
		(0.050)	(0.035)***	(0.044)*	(0.010)	(0.029)	(0.184)
1(Spiff and 100 Rebate)		0.063	0.172	0.107	-0.033	0.155	0.289
		(0.067)	(0.064)***	(0.066)	(0.016)**	(0.065)**	(0.259)
R^2	0.39	0.17	0.22	0.14	0.20	0.25	0.16
N	20,240	2,068	2,068	2,068	2,067	1,742	2,068
Dep. Var. Control Mean	0	.13	0	.05	0	.01	.41

Notes: This table reports the estimates of Equation (3.9). All regressions include agent and phase indicator variables. Robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively.

the website directed them to. Columns 4 and 5 show that agents disclosed elements of the information script (quoted a payback period and/or said that Energy Star saves money) on 9 to 17 percent more information treatment calls relative to control.

Column 6 reports whether an agent said something approximating one of the treatment scripts during the call. These results are consistent with the results in columns 3, 4, and 5, in that agents appear to comply with the experiment protocol on about 15 to 20 percent of calls. The number of observations is lower in this column because we did not begin to record this variable until after the first 326 audits were completed. The outcome variable in column 7 is the natural log of one plus the estimated number of seconds that the agent and customer discussed energy efficiency; it is 33 to 62 log points larger in the information and rebate treatment groups.

The bottom row of Table 3.4 gives the mean of each dependent variable in the control group. Agent-reported compliance in column 1 is zero by definition in the control group because the website did not ask agents to report whether they complied on control group calls. Although agents do mention Energy Star on some control group calls, they almost never deliver a script erroneously to the control group: out of the more than 400 control group calls that were audited, agents quoted a payback period once and mentioned an Energy Star rebate twice. In total, they gave information that sounded like one of the treatment scripts to the control group less than one percent of the time. Column 3 shows that agents did mention a rebate on a small but statistically significant share of non-rebate treatment group calls - both Information Only and Spiff Only calls. This may reflect some small amount of recording error in the audits or mistakes by the sales agents.

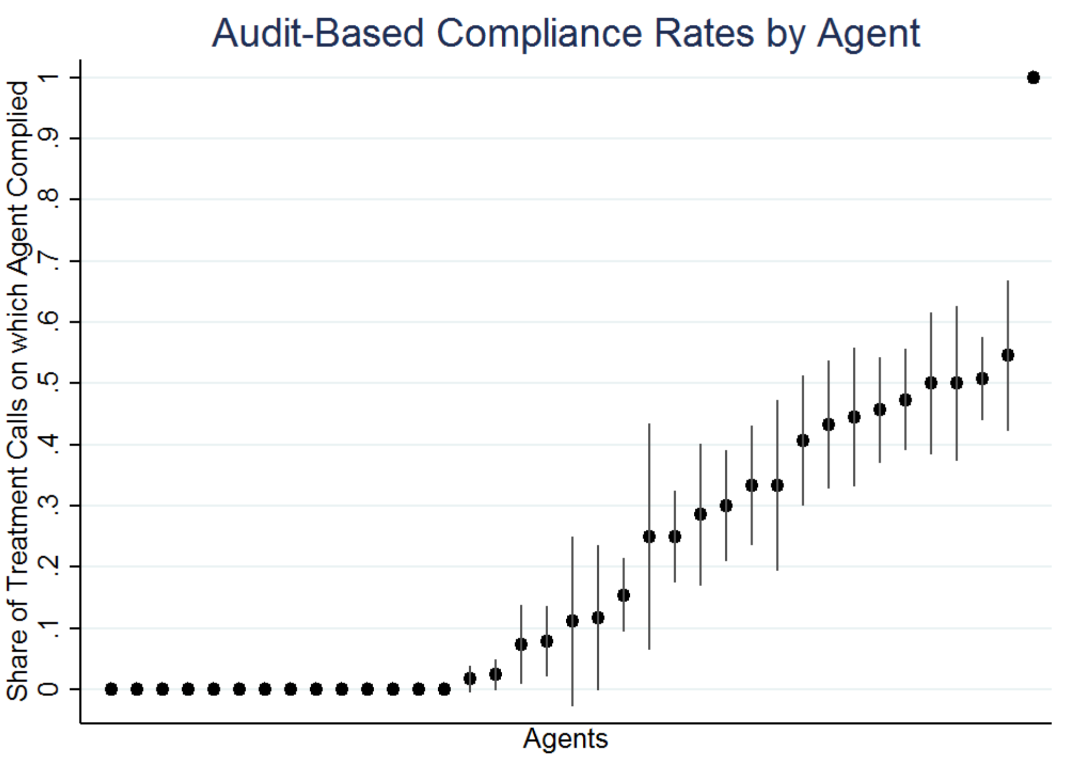
3.5.1.1 Measuring Sales Agent Compliance

For use in the next section, we use the audit data to construct a measure of compliance with the experiment scripts. This variable is intended to measure explicit compliance with the experiment scripts, not any other form of discussion of energy efficiency. For consumers whose calls were audited, we define an indicator variable N^+_{iat} that takes value 1 if an agent read a script on any treatment group call, mentioned an Energy Star rebate on a rebate call, or quoted a payback period on an information call. N^+_{iat} takes value zero otherwise.

There is substantial variation in compliance across agents. Define $\overline{N}^+_a$ as the mean of N^+_{iat}

across all of agent a's calls, excluding the Spiff Only and control calls. Figure 3.1 shows the CDF of $\overline{N}_a^+$ for all agents who were audited more than five times. About one-quarter of agents never comply, the median $\overline{N}_a^+$ is 11 percent, and one quarter of agents comply more than 40 percent of the time. This dispersion implies that we can exploit variation in compliance rates across agents to improve power in tests of the effects of information provision on demand. Using the data in this graph, we group agents into four compliance groups with $\overline{N}_a^+ = 0$, $0 < \overline{N}_a^+ \leq 0.2$, $0.2 < \overline{N}_a^+ \leq 0.4$, and $\overline{N}_a^+ \geq 0.4$. We define G_a as the mean of $\overline{N}_a^+$ across all agents within agent a's compliance group, where agent a is the agent with whom consumer i had his or her final call. The mean values of G_a for agents in the four compliance groups are 0, 0.08, 0.29, and 0.53 respectively. Agents who were audited fewer than five times are automatically assigned the G_a for the second compliance group, which includes the median $\overline{N}_a^+$.

Figure 3.1: CDF of Audited Compliance



Notes: This figure plots the average compliance rate on all calls other than Spiff Only and control group calls, for all agents that were audited more than five times. Compliance is measured by an indicator variable N_{ia}^+ that takes value 1 if an agent read a script, mentioned an Energy Star rebate on a rebate call, or quoted a payback period on an information call.

3.5.1.2 Spillovers of Information Provision to Non-Information Group Calls

One reason why agents might not provide much information about Energy Star is that they might not know what to say, or might not be well-practiced in discussing energy efficiency. To test whether agents learn to disclose information, we exploit the fact that the experiment induced Information Treatment Agents to repeatedly deliver the Energy Star informational script, while Information Control Agents were never directly exposed. We regress the same compliance measures from Table 3.4 on the interaction of Information Treatment Agent indicator variables with a vector of treatment group indicators. Defining I_a as an indicator variable for whether agent a is assigned as an Information Treatment Agent, the regression is:

$$N_{iat} = \gamma \mathbf{T}_i I_a + \beta \mathbf{T}_i + \phi_t + \upsilon_{iat} \tag{3.10}$$

Standard errors are clustered by agent. Table 3.5 presents the estimated γ coefficients on the interactions of I_a with call treatment assignment indicators $\mathbf{T}_i$. The table parallels Table 3.4, with three exceptions. First, to increase power, we combine the \$25 and \$100 Rebate groups into one indicator. Second, the samples exclude information treatment calls, because Information Control Agents do not have any information treatment calls, and the objective is to compare information provision on non-information group calls. Third, we do not present regressions for *Payback Period*, because agents only quoted payback periods on four audited non-information group calls. (All four involved Information Treatment Agents.)

Table 3.5 shows that Information Treatment Agents are not more likely than Information Control Agents to mention experimental rebates on calls in any treatment group. Information Treatment Agents are, however, more likely to mention Energy Star on control calls and to mention that Energy Star saves money on Spiff Only calls. Column 6 shows that Information Treatment Agents talk about Energy Star for approximately 30 percent longer on Spiff Only and control group calls. The standard errors are too wide to determine whether there is a meaningful difference in information provision on combination Spiff plus Rebate calls.

Table 3.5: Information Treatment Agents vs. Information Control Agents

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Agent Reported	Mentioned E-Star	Rebate	Saves Money	Read Script	ln(E-Star Seconds)
1(Rebate)	-0.072	0.063	-0.048	0.045	0.005	0.189
	(0.098)	(0.102)	(0.085)	(0.029)	(0.100)	(0.339)
1(Spiff Only)		0.092	0.004	0.062	0.030	0.314
		(0.059)	(0.035)	(0.031)*	(0.023)	(0.180)*
1(Spiff and Rebate)		-0.143	-0.136	-0.076	-0.090	-0.555
		(0.144)	(0.109)	(0.122)	(0.084)	(0.532)
1(Control)		0.109	0.003	0.044	0.007	0.320
		(0.052)**	(0.007)	(0.029)	(0.010)	(0.157)**
R^2	0.01	0.04	0.09	0.02	0.11	0.04
N	8,276	1,642	1,642	1,642	1,408	1,642
Dep. Var. Control Agent Mean	.527	.178	.101	.051	.101	.582

Notes: This table reports the estimates of γ in Equation (3.10). All regressions include phase and treatment group indicator variables. Robust standard errors, clustered by agent, in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively.

These results have two implications. First, they suggest that one reason why the Retailer's sales agents do not frequently discuss information is that they were not well-practiced at doing so. Once Information Treatment Agents learned how to discuss Energy Star on information treatment calls, they began to do so without explicit instruction on control and spiff calls. In the context of our model, the experiment reduced the cost of information provision α for Information Treatment Agents. Second, these results imply that the estimates of effects of information disclosure on sales should take account of spillovers, in the form of increased discussion of energy efficiency on control group calls.

3.5.2 Effects on Consumer Choice

We examine two binary outcomes Y_{iat}: whether the consumer purchases any model from the Retailer and whether the consumer purchases an Energy Star model. For each outcome, we run three specifications. The first specification is the intent-to-treat (ITT) estimator:

$$Y_{iat} = \tau \mathbf{T}_i + \phi_i + \mu_a + \varepsilon_{iat} \tag{3.11}$$

As above, this includes agent and phase indicator variables.[11]

The second specification is an instrumental variables (IV) estimator, substituting agent-reported compliance N_{iat}^s for $\mathbf{T}_i$ in Equation (3.11) and instrumenting for N_{iat}^s with $\mathbf{T}_i$.[12] This sample is smaller because it excludes the spiff treatment calls, as N_{iat}^s is undefined for these calls. The third specification is what we call the "Scaled ITT" estimate: we interact $\mathbf{T}_i$ with G_a, which reflects the probability that agent a delivered the specific information or rebate script to consumer i. Intuitively, multiplying by compliance probability scales the τ coefficient to be equivalent to a Local Average Treatment Effect (LATE).[13]

In this context, the ITT and LATE are likely to bound the average treatment effect of providing information to all consumers. While the ATE could theoretically be larger than the LATE if sales agents targeted information at the least responsive consumers, we shall see momentarily that agents appear to target consumers who are more interested in Energy Star. If all treatment group consumers with $N_{iat} = 0$ would have had zero treatment effect, then the ITT equals the ATE. On the other hand, if the agents quasi-randomly chose whom to disclose to, then consumers with $N_{iat} = 0$ and $N_{iat} = 1$ would have the same treatment effect, and the LATE would equal the ATE.

Table 3.6 presents the results. The left three columns show effects on overall sales of any model from the Retailer. The right three columns show effects on Energy Star sales. Within each set of three columns, the first is the ITT, the second is the IV using agent-reported compliance, and the third regression is the Scaled ITT. Because the interaction effects between information and the two rebate level are never statistically significant, we drop these terms and report results for six major treatment groups relative to control. At the bottom of the table, we report the mean purchase probabilities in the control group: about 36 percent of consumers purchase from the Retailer, and about 0.9 percent of consumers purchase an Energy Star model.

[11] Once we control for phases, additional time controls do not improve consistency or efficiency. Adding month-of-sample indicators, for example, does not change coefficients or standard errors. Furthermore, the estimates do not change when we exclude Phase I, which pre-dated the introduction of the spiff.

[12] Our previous working paper contained an error in this sentence that incorrectly described our IV procedure.

[13] N_{iat}^+ is not statistically different across phases or treatment groups, so we cannot increase precision by also projecting ϕ_i or $\mathbf{T}_i$ onto G_{ia}. Although G_{ia} is a mean calculated with sampling error using the audit data, we calculate that adjusting for sampling error in this generated regressor has only a small impact on the standard errors.

Table 3.6: Treatment Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	1(Sale)	1(Sale)	1(Sale)	1(EStar)	1(EStar)	1(EStar)
1(100 Rebate)	-0.010	-0.024	0.030	0.006	0.012	0.037
	(0.011)	(0.022)	(0.045)	(0.003)**	(0.005)**	(0.013)***
1(25 Rebate)	-0.019	-0.040	-0.015	0.001	0.002	0.003
	(0.007)***	(0.015)***	(0.029)	(0.001)	(0.003)	(0.005)
1(Information)	-0.005	-0.009	-0.059	0.000	0.000	0.004
	(0.009)	(0.020)	(0.036)	(0.002)	(0.004)	(0.007)
1(Spiff)	0.007		0.025	-0.002		0.001
	(0.012)		(0.042)	(0.002)		(0.007)
1(Spiff and 25 Rebate)	0.000		0.055	-0.004		-0.007
	(0.032)		(0.111)	(0.002)**		(0.005)
1(Spiff and 100 Rebate)	0.041		0.200	0.040		0.219
	(0.053)		(0.193)	(0.022)*		(0.118)*
R^2	0.02	0.02	0.02	0.01	0.01	0.01
N	23,347	20,240	23,347	23,347	20,240	23,347
Dep. Var. Control Mean	.364	.364	.364	.009	.009	.009
Regression Type:	ITT	Self-Report IV	Scaled ITT	ITT	Self-Report IV	Scaled ITT

Notes: This table reports the estimates of Equation (3.11). All regressions contain agent and period indicator variables. Robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively.

The treatments have no effect on overall sales, except that the ITT and self-report IV suggest that a $25 rebate may reduce purchase probability. One explanation for this is that even mentioning a small rebate for a different model generates a version of choice overload, complicating the sales interaction and causing a slight decrease in purchase probability. This would be consistent with other evidence of choice overload, such as Iyengar and Lepper (2000). Another explanation is that the result is spurious and would not replicate. We find this latter explanation more plausible, partially because there is no negative effect in the Scaled ITT. Intuitively, the Scaled ITT differs from the ITT because it weights more heavily the treatment effects from more compliant agents. The fact that the negative association disappears in the Scaled ITT implies that the ITT effect is driven by agents who aren't actually doing anything to comply with the experiment, which suggests a spurious correlation. The first two columns of Appendix Table A.2 replicate column 1 for agents in the bottom two vs. top two compliance groups, confirming that none of the treatments affect sales in the subsample of more compliant agents.

Although there appears to be little or no effect on the Retailer's overall sales, columns 4-6 show that the treatments do shift the composition of sales toward Energy Star. The $100 rebate increases Energy Star purchase probability by 0.006 to 0.037 percentage points in the three specifications.[14] This effect is large against a control group market share of 0.9 percent, although it is small as a share of the total potential market. The combination of a spiff and $100 rebate has a very large point estimate, increasing purchase probability by 4 to 22 percentage points. The standard errors are wide, as this treatment was only offered to a small share of consumers in phase 3 of the experiment.

The information treatments have a tightly estimated zero statistical effect. In column 6, the standard errors are tight enough to bound the Local Average Treatment Effect at less than 1.5 percentage points. While this would represent a large percent increase in Energy Star purchase probability given the small base, the effect in percentage point terms is economically small.

Earlier, we documented that although agents effectively never explicitly read a treatment script on a control call, Information Treatment Agents are more likely to mention Energy Star on control calls. If this increases the probability of Energy Star purchase, then the estimates in Table 3.6 would understate the true effects of the information scripts. The audit data allow us to address such spillovers. We construct an alternative "Scaled ITT" estimator, interacting $\mathbf{T}_i$ with each compliance group's average difference in *Mentioned E-Star* between treatment and control calls. Intuitively, this scales the treatment effect to equal a LATE with *Mentioned E-Star* as the endogenous variable. The third column of Appendix Table A.2 presents the results. The coefficients and standard errors are inflated, as one should expect from the fact that the coefficients in Table 3.4 are smaller for *Mentioned Energy Star* than for the variables used to construct G_{ia}. The qualitative results are similar: the Spiff and $100 Rebate combination still has a very large effect, and the standard errors bound the information effect at no more than 2.7 percent with 90 percent confidence.

As discussed in Section 3.2.3, Energy Star models are only substitutable with some standard

[14] In column 4, the combination of the spiff and a $25 rebate appears to reduce demand for Energy Star. We suspect that this result also would not replicate, and it is not statistically significant in the Scaled ITT in column 6.

models. Under the assumption that the treatments do not affect whether or not consumers purchase a substitutable model, we can consistently estimate treatment effects within the sample to the set of consumers that purchase substitutable models. Appendix Table A.3 presents results. The coefficients are larger than in Table 3.6, as one should expect from excluding consumers with smaller treatment effects.[15] The qualitative results are also similar: the Spiff and $100 Rebate combination has a very large effect, and the standard errors bound the information effect at no more than 4.9 percentage points with 90 percent confidence in the "Scaled ITT." For comparison, Energy Star represents 3.4 percent of substitutable models. The standard errors suggest that even if sales agents provided information to all callers, Energy Star would still not represent more than 4.9+3.4=8.3 percent of the market of substitutable models. Thus, the lack of seller-provided information does not explain much of the low takeup of energy efficient water heaters.

Appendix Table A.4 presents alternative estimates using a probit estimator; the signs and discrete significance levels are the same or stronger.

3.5.3 Targeted Information Disclosure

Table 3.4 documents that the sales agents only partially comply with the experiment. Are agents strategic in providing information to consumers that are more interested in energy efficiency? Recall that N^s_{iat} is an indicator for whether the consumer is in a treatment group and the agent reported compliance with delivering the script, and define T_i as an indicator for whether consumer i is in any rebate or informational treatment group. We exclude the spiff treatment calls because there was no script for the agent to "comply" with on these calls. Table 3.7 reports estimates of the following regression:

$$Y_{iat} = \xi T_i N^s_{iat} + \kappa T_i (1 - N^s_{iat}) + \phi_i + \mu_a + \varepsilon_{iat} \tag{3.12}$$

As outcome variables Y_{iat}, we use two different measures of interest in energy efficiency, both of which could be affected by the treatments. To construct the dependent variable in

[15]Energy Star model availability and consumer preferences vary by geography, so if geography were somehow imbalanced across treatment groups, our coefficient estimates would be biased. Appendix Table A.3 also shows that the coefficient estimates are very similar when also including indicators for each purchaser's three-digit zip code. Of course, this is to be expected in a randomized experiment.

column 1, we exploit an open-answer question from the follow-up survey: *What were the two most important factors in your water heater purchase decision?* The dependent variable is an indicator taking value one for the consumers who had one of their two factors coded as "saving energy and/or environmental conservation." The dependent variable in columns 2-4 is an indicator for whether the consumer purchased Energy Star. These regressions use agent-reported compliance because the sample of audits is too small for sufficient power.

Table 3.7: Targeted Information Provision

	(1)	(2)	(3)	(4)
Dependent Variable:	1(Factor)	1(EStar)	1(EStar)	1(EStar)
T x Agent Reported Compliance	-0.014	0.013	0.038	0.034
	(0.030)	(0.002)***	(0.007)***	(0.006)***
T x (1 - Agent Reported Compliance)	-0.057	-0.009	-0.031	-0.022
	(0.021)***	(0.001)***	(0.005)***	(0.004)***
R^2	0.18	0.02	0.05	0.05
N	404	20,240	5,180	6,123
Dep. Var. Control Mean	.061	.009	.033	.025

Notes: This table reports the estimates of Equation (3.12). All regressions contain agent and period indicator variables. Robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively.

The coefficient κ measures the difference in Y_{iat} between treatment group consumers who did not receive the treatment and the control group, while ξ measures the difference between treatment group consumers who did receive the treatment (according to the agents' self-reports) and the control. ξ is a mix of selection effects and treatment effects, while κ is purely selection effects. If $\kappa < 0$, this implies that the consumers to whom agents are not disclosing information are less likely than average to purchase Energy Star, and thus that consumers who are provided with information are more likely than average to purchase Energy Star. While this is not the same as targeting information disclosure at consumers who will have the largest treatment effects, it implies targeting in a different and likely related sense.

In all four columns of Table 3.7, the statistically negative estimates of κ shows that agents are more likely to report delivering the script to consumers that are more interested in energy efficiency. Column 3 includes in the sample only consumers who purchase substitutable models. Even within this group, agents still target consumers with a higher probability of purchasing

Energy Star. Column 4 includes controls for the nine address-based demographic variables. Because these covariates are missing for consumers who did not purchase from the Retailer, the sample is limited to consumers who purchased from the Retailer. The fact that κ is still negative after conditioning on observables shows that agents target information provision based on other unobservable factors.

How is such targeting possible? Based on conversations with sales agents and with our research assistant who carried out the audits, we believe that agents learn about the consumer's preferences as the call continues. Agents may have delayed reading the Energy Star script until later in the call, after having the chance to gauge the consumer's receptiveness.

3.5.4 Relationship to the Model

Our main empirical results tell a coherent story that is consistent with the three propositions in Section 3.4. Proposition 1 states that agents will not inform consumers if information does not increase demand. Results in Table 3.6 and related appendix tables show that information increases Energy Star sales by at most a few percentage points, even when focusing on the Scaled ITT estimates and limiting the sample only to consumers who are purchasing water heaters that are substitutable with Energy Star models. Consistent with this and with Proposition 1, the audit data show that agents spent very little time discussing Energy Star unless the experiment website directed them to. Agents mentioned Energy Star on only 13 percent of control group calls, and the median time spent discussing Energy Star within this 13 percent was 15 seconds.

Proposition 2 states that agents will provide more information when incentivized to do so, but they will not be very responsive to sales incentives if information has small effects on demand. The audit data show that agents provide at best slightly more information about Energy Star on no-rebate spiff calls compared to control: in Table 3.4, the coefficient on 1(Spiff Only) is positive but not statistically significant in columns 2, 4, 6, and 7.

Both of Propositions 1 and 2 illustrate the interaction between the supply and demand sides of the market. It was difficult for the Retailer to increase information provision through the experiment because their sales agents knew that most consumers would be unresponsive to the information. The small Local Average Treatment Effects in the IV and Scaled ITT estimators, combined with the targeting results in Table 3.7, suggest that agents' lack of compliance is best characterized as "strategic," not "shirking." In the context of the model, we mean that agents' information disclosure cost α is low enough that if more consumers were "High types" that would potentially be interested in Energy Star, agents might devote more effort to marketing those models.

Proposition 3 states that sales incentives and customer rebates reinforce each other, because the rebate increases sales agents' gains from responding to the sales incentive. Although the standard errors are wide, the point estimates in the sales data in Table 3.6 strongly support this: the combined Spiff and $100 rebate treatment substantially increases Energy Star purchase probability. The audit data provide evidence that is suggestive of the microfoundations for this result. Specifically, agents appear to be slightly more likely to market Energy Star on the spiff treatment calls as the rebates increase from $0 to $25 to $100: in Table 3.6, the point estimates in columns 2, 3, 4, 6, and 7 all increase as the rebates increase.

Finally, the model considers a potential role for government-provided sales incentives, which could perhaps be a complement to product subsidies traditionally offered by governments and utility energy efficiency programs. The empirical results in Table 3.6 suggest a strong complementarity between price reductions and sales incentives. The model clarifies that government-provided incentives could increase welfare in the presence of an informational market failure, as long as firms are not able to fully "undo" the government incentive by changing their agents' compensation structure.

3.5.5 Why Is Demand for Energy Star Low?

Equation (3.7) shows that a minimum energy efficiency standard is more likely to increase welfare if the energy efficient product generates utility gains for a larger share of the population or if the cost of providing information is high. Our results show that the experimental attempts at information provision do not significantly increase Energy Star demand. Does this suggest that minimum energy efficiency standards might increase welfare by addressing imperfect information?

A first potential explanation for our empirical results is that although the Energy Star product would generate gains for many consumers, the Retailer is not able to credibly inform consumers of this, perhaps due to time constraints on the sales interactions or an inability to credibly convey attributes of this higher-priced product. If this were true, many consumers would remain unaware of Energy Star and unconvinced of the cost savings. A second explanation is that the Energy Star model would not in fact generate utility gains for many consumers, and consumers make an informed decision to not purchase the product. If this were true, consumers would be aware of the Energy Star product and the potential energy cost savings.

We use results from the follow-up surveys to measure the importance of these two explanations. Near the end of the survey, we asked consumers, *Some water heater models that use less energy are officially designated as Energy Star. Did you buy an Energy Star model?* The results suggest substantial confusion, consistent with the first explanation. The top panel of Table 3.8 shows that while only 2.1 percent actually purchased Energy Star, 52 percent of survey respondents think that they did. This should be interpreted cautiously, as there could have been experimenter demand effects: respondents may not have wanted to tell the interviewer that they had not purchased the Energy Star option. We designed the survey specifically to reduce demand effects: this was the first question we asked about energy efficiency, and we asked it after a series of other questions that signaled that the interviewer was not particularly interested in energy.

Table 3.8: Survey Results: Energy Star Purchases

Self-Reported Energy Star Purchase	
Some water heater models that use less energy are officially designated as Energy Star. Did you buy an Energy Star model?	
Response	**Percent**
Yes	52
No	24
Not sure	24
True Energy Star share in survey sample	2.1

Reasons to Not Purchase Energy Star	
Why did you decide to buy a standard, non-Energy Star model over an Energy Star model?	
Response	**Percent**
Upfront price too high	33.3
I was not aware that there was an Energy Star option	15.4
Energy Star not in stock	8
Needed a short tank	5.9
No electrical outlet	2.8
Needed low-NOx	2.3
Wanted longer warranty	1.4
Non-Energy Star heats water faster	1.2
Other	30.7

This table presents responses to two questions from our follow-up survey of consumers who purchased from the Retailer. Sample size for the first question is 891. The second question was asked only of people who thought they had not purchased an Energy Star model or were not sure but in fact had not; sample size is 423.

Respondents who said they had not purchased Energy Star, or who were "Not Sure" but in fact had not, were asked why they had not purchased Energy Star. The bottom panel of Table 3.8 shows that the primary reason was high prices. This is consistent with the second explanation of informed consumers choosing not to purchase Energy Star. Fifteen percent of these consumers (or about seven percent of the entire surveyed population) report being unaware that there was an Energy Star option.

The follow-up survey also elicited beliefs over energy costs for standard models and energy cost savings from Energy Star models. Table 3.9 shows the 10th, 50th, and 90th percentiles of beliefs, along with the mean and the best estimate of the true value, from the yellow tags. The first question in Table 3.9 shows that consumers' mean beliefs about water heater energy costs are approximately in line with the yellow tags. The second question, however, suggests that the mean and median consumers actually overestimate the average dollar value of potential cost savings. The third question takes the ratio of the second question to the first, which translates to a percent savings and can account for heterogeneity in consumers' utilization. The "true" mean energy cost savings on the yellow tags are about 10 percent, while the median and mean consumers report believing that Energy Star could save 25 and 32 percent, respectively.[16] Of course, the beliefs were elicited in a phone survey and not made incentive compatible, so they should be interpreted cautiously. Notwithstanding, they suggest that underestimating energy cost savings is not a barrier to Energy Star takeup.

Table 3.9: Survey Results: Beliefs About Energy Star

How much money do you think the natural gas for the water heater will cost each year?				
10th	**50th**	**90th**	**Mean**	**Yellow Tag**
50	200	600	305	Approx 300
How much less money do you think the natural gas would cost each year for an Energy Star water heater compared to a similarly-sized non-Energy Star water heater?				
10th	**50th**	**90th**	**Mean**	**Yellow Tag**
0	50	300	129	Approx 30
Implied percent savings from Energy Star				
10th	**50th**	**90th**	**Mean**	**Yellow Tag**
5	25	67	32	Approx 10

This table presents responses to three questions from our follow-up survey of consumers who purchased from the Retailer. Sample size is 891.

3.6 Conclusion

Imperfect information is an oft-cited reason why regulators intervene in markets for energy-using durables though mandatory information disclosure, subsidies, and standards. In theory, one natural way for consumers to learn about the benefits of energy efficient products is through retailers. Motivated by this, we partnered with a large nationwide retailer to test the effects of information provision, customer rebates, and sales incentives on the behavior of both sales agents and consumers. We also present a simple model which, while not a theoretical advance *per se*, helps to motivate the experiment and clarify the results.

Results show that retailer-provided information is ineffective at increasing Energy Star demand, even after adjusting for partial compliance by sales agents. Knowing that information would be ineffective, sales agents appear to have marketed Energy Star to only the most receptive consumers, strategically failing to provide information to the majority. Follow-up surveys

[16] Additional (unreported) regressions show that confusion about Energy Star purchases and beliefs about Energy Star savings do not vary across treatment groups, although estimates are imprecise.

provide some evidence that at least some consumers who bought from the Retailer are unaware of the Energy Star product or confused about what they have bought. The majority of consumers who bought from the Retailer and completed our survey, however, are aware of Energy Star and may even overestimate its benefits. These results highlight the difficulties that retailers can face in increasing demand for energy efficient or otherwise “high-quality” products. In this context, a key difficulty appears to be that many consumers still don’t view energy efficiency as a privately-beneficial investment, even after the retailer’s attempts to inform them.

The water heater market has some unusual features, and our sample comprises only a small share of that market. What broader lessons can be drawn from our results? First, our results fit into a broader set of results from the literature suggesting that while social comparisons or various forms of persuasion can motivate pro-environmental behaviors, hard information about energy costs seems to have more limited effects. As our model makes clear, “hard” information about product availability (as well as product costs and benefits) is more relevant from a welfare perspective because a regulator’s inability to address imperfect information could justify corrective policies, while the welfare implications of persuasion are less clear. Our experimental design does not allow us to cleanly distinguish whether consumers are more generally uninformed or whether they are simply disinterested in information from the retailer’s agents. Notwithstanding, our survey results help somewhat to resolve this question. Furthermore, the water heater market is certainly one where consumers begin the process with very little information and do rely on agents’ knowledge for many other aspects of the purchase.

Second, our theoretical and empirical results clearly suggest that policymakers interested in maximizing the effects of subsidy dollars on Energy Star purchases might consider allocating some incentives to sales agents, not just to consumers. While the magnitude of our effect size is certainly specific to our context, this general suggestion is novel. Third, our results provide unusually granular insight into the role of sales agents in the process of information disclosure. In many situations like ours, a firm or policymaker must consider how to incentivize agents if they want to increase information disclosure. In equilibrium, firms and policymakers can also potentially learn about the value of information from whether experienced sales agents choose to provide it.

References

[1] Ackerberg, Daniel, Lanier Benkard, Steve Berry, and Ariel Pakes (2007). “Econometric Tools for Analyzing Market Outcomes.” *Handbook of Econometrics*, Vol. 6A.

[2] Allcott, Hunt (2011). “Social Norms and Energy Conservation.” *Journal of Public Economics*, Vol. 95, No. 9-10 (October), pages 1082-1095.

[3] Allcott, Hunt (2013). “The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market.” *American Economic Journal: Economic Policy*, Vol. 5, No. 3 (August), pages 30-66.

[4] Allcott, Hunt, and Todd Rogers (2014). “The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation.” *American Economic Review*, forthcoming.

[5] Allcott, Hunt, and Dmitry Taubinsky (2015). “Evaluating Behaviorally-Motivated Policy: Experimental Results from the Lightbulb Market.” *American Economic Review*, forthcoming.

[6] Anagol, Santosh, Shawn Cole, and Shayak Sarkar. ”Understanding the Advice of Commissions-Motivated Agents: Evidence from the Indian Life Insurance Market.” Harvard Business School Working Paper, No. 12-055, (March 2013.)

[7] Anderson, Dennis, and John Claxton (1982). ”Barriers to Consumer Choice of Energy Efficient Products.” *Journal of Consumer Research*, Vol. 9, No. 2 (September), pages 163-170.

[8] Andersen, Soren and Andrew Elzinga (2014). “A Ban on One is a Boon for the Other: Strict Gasoline Content Rules and Implicit Ethanol Blending Mandates.” *Journal of Environmental Economics and Management* 67 (3): pages 258-273.

[9] Andrews, Donald W.K. and Gustavo Soares (2010). “Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection.” *Econometrica*, 78(1): 119-157.

[10] Angrist, Joshua, and Jorn-Steffen Pischke (2012). “Probit better than LPM?” Available from http://www.mostlyharmlesseconometrics.com/2012/07/probit-better-than-lpm/

[11] Ashraf, Nava, Oriana Bandiera, and Kelsey Jack (2012). “No Margin, No Mission? A Field Experiment on Incentives for Pro-Social Tasks.” SSRN Working Paper No. 2013825 (February).

[12] Auffhammer, Maximilian and Ryan Kellogg (2011). “Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality.” *American Economic Review*, 101(4), pages 2687-2722.

[13] Bajari, Pat, Lanier Benkard and Jonathan Levin (2007). “Estimating Dynamic Models of Imperfect Competition.” *Econometrica*, p 1331-1370.

[14] Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2005). "Social Preferences and the Response to Incentives: Evidence from Personnel Data." *Quarterly Journal of Economics*, Vol. 120, No. 3 (August), pages 917-962.

[15] Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2009). "Social Connections and Incentives in the Workplace: Evidence from Personnel Data." *Econometrica*, Vol. 77, No. 4, pages 1047-1094.

[16] Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2010). "Social Incentives in the Workplace." *Review of Economic Studies*, Vol. 77, No. 2 (April), pages 417-458.

[17] Bandiera, Oriana, and Imran Rasul (2006). "Social Networks and Technology Adoption in Northern Mozambique." *The Economic Journal*, Vol. 116, No. 514, pages 869-902.

[18] Barankay, Iwan (2012). "Rank Incentives: Evidence from a Randomized Workplace Experiment." Working Paper, Wharton (July).

[19] Becker, Gary, and Kevin Murphy (1993). "A Simple Theory of Advertising as a Good or Bad." *Quarterly Journal of Economics*, Vol. 108, No. 4 (November), pages 941-964.

[20] Becker, Randy A. (2005). "Air Pollution Abatement Costs Under the Clean Air Act: Evidence from the PACE Survey." *Journal of Environmental Economics and Management*, 50(1), 144-169.

[21] Berman, Eli and Linda Bui (2001). "Environmental Regulation and Productivity: Evidence from Oil Refineries." *Review of Economics and Statistics*, 83(3): 498–510.

[22] Berry, Steven, James Levinsohn, and Ariel Pakes (1995). "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841–890.

[23] Berry, Steven and Peter Reiss (2007). "Empirical Models of Entry and Market Structure." *Handbook of Econometrics*, Vol. 6A.

[24] Bhargava, Saurabh, and Dayanand Manoli (2013). "Why Are Benefits Left on the Table? Assessing the Role of Information, Complexity, and Stigma on Take-up with an IRS Field Experiment." Working Paper, University of Texas at Austin.

[25] BLS (Bureau of Labor Statistics) (2014). "Consumer Expenditure Survey: 2011 Expenditure Tables." Available from http://www.bls.gov/cex/csxstnd.htm

[26] Bollinger, Brian, Phillip Leslie, and Alan Sorensen (2011). "Calorie Posting in Chain Restaurants." *American Economic Journal: Economic Policy*, Vol. 3, No. 1 (February), pages 91-128.

[27] Borenstein, Severin and Andrea Shepard (2002). "Sticky Prices, Inventories, and Market Power in Wholesale Gasoline Markets." *RAND Journal of Economics*, Vol. 33, No 1.

[28] Brandon, Alec, John List, Robert Metcalfe, Michael Price (2014). "What Dives the Adoption of Energy Efficient Technologies? The Role of Social Information and Advertisements." Working Paper, University of Chicago (April).

[29] Brown, Jennifer, Justine Hastings, Erin T. Mansur, and Sofia B. Villas-Boas (2008). "Reformulating Competition? Gasoline Content Regulation and Wholesale Gasoline Prices." *Journal of Environmental Economics and Management*, Vol. 55, No. 1, pages 1-19.

[30] Bulow, Jeremy, John Geanakoplos, and Paul Klemperer (1985). "Multimarket Oligopoly: Strategic Substitutes and Complements." *Journal of Political Economy*, Vol. 93, no. 3.

[31] Chakravorty, Ujjayant, Celine Nauges, Alban Thomas (2008). "Clean Air Regulation and Heterogeneity in US Gasoline Prices." *Journal of Environmental Economics and Management*, Vol. 55, pages 106-122.

[32] Chan, Tat, Jia Li, and Lamar Pierce (2014). "Compensation and Peer Effects in Competing Sales Teams." *Management Science*, forthcoming.

[33] Chatterji, Aaron, Michael Findley, Nathan Jensen, Stephan Meier, and Daniel Nielson (2015). "Field Experiments in Strategy Research." *Strategic Management Journal*, forthcoming.

[34] Chesnes, Matthew (2014). "The Impact of Outages on Prices and Investment in the US Oil Refining Industry." *(working paper)*

[35] Choi, James, David Laibson, and Brigitte Madrian (2010). "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds." *Review of Financial Studies*, Vol. 23, No. 4, pages 1405-1432.

[36] Chouinard, Haley H. and Jeffrey M. Perloff, "Gasoline Price Differences: Taxes, Pollution Regulations, Mergers, Market Power and Market Conditions", *The B.E. Journal of Economic Analysis and Policy* , *7*(1), article 8.

[37] Davis, Lucas and Lutz Killian (2001). "Estimating the Effect of a Gasoline Tax on Carbon Emissions." *Journal of Applied Econometrics*, 16(7), 1187-1214.

[38] Davis, Lucas, and Gilbert Metcalf (2014). "Does Better Information Lead to Better Choices? Evidence from Energy-Efficiency Labels." NBER Working Paper No. 20720 (November).

[39] Deutsch, Matthias (2010a). "Life Cycle Cost Disclosure, Consumer Behavior, and Business Implications: Evidence from an Online Field Experiment." *Journal of Industrial Ecology*, Vol. 14, No. 1, pages 103-120.

[40] Deutsch, Matthias (2010b). "The Effect of Life-Cycle Cost Disclosure on Consumer Behavior: Evidence from a Field Experiment with Cooling Appliances." *Energy Efficiency*, Vol. 3, pages 303-315.

[41] DEHWA (Australian Department of the Environment, Water, Heritage, and the Arts) (2008). "Regulatory Impact Statement: Proposal to Phase-Out Inefficient Incandescent Light Bulbs." http://www.energyrating.gov.au/wp-content/uploads/Energy_Rating_Documents/Library/Lighting/Incandescent_Lamps/200808-ris-phaseout.pdf

[42] Dixit, Avinash, and Victor Norman (1978). "Advertising and Welfare." *Bell Journal of Economics*, Vol. 91, No. 1 (April), pages 1-17.

[43] DOE (U.S. Department of Energy) (2010). Energy Star Water Heater Market Profile." Available from https://www.energystar.gov/ia/partners/prod_development/new_specs/downloads/water_heaters/Water_Heater_Market_Profile_2010.pdf

[44] Dolan, Paul, and Robert Metcalfe (2013). "Neighbors, Knowledge, and Nuggets: Two Natural Field Experiments on the Role of Incentives on Energy Conservation." CEP Discussion Paper No. 1222 (June).

[45] Dorfman, Robert, and Peter Steiner (1954). "Optimal Advertising and Optimal Quality." *American Economic Review*, Vol. 44, No. 5 (December), pages 826-836.

[46] Dranove, David, and Ginger Zhe Jin (2010). "Quality Disclosure and Certification: Theory and Practice." *Journal of Economic Literature*, Vol. 48, No. 4 (December), pages 935-963.

[47] Duarte, Fabian, and Justine Hastings (2012). "Fettered Consumers and Sophisticated Firms: Evidence from Mexico's Privatized Social Security Market." NBER Working Paper 18582 (November).

[48] Duflo, Esther, and Emmanuel Saez (2003). "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *Quarterly Journal of Economics*, Vol. 118, No. 3 (August), pages 815-842.

[49] Duflo, Esther, William Gale, Jeffrey Liebman, Peter Orszag, and Emmanuel Saez (2006). "Saving Incentives for Low- and Middle-Income Families: Evidence from a Field Experiment with H&R Block." *Quarterly Journal of Economics*, Vol. 121, No. 4 (November), pages 1311-1346.

[50] Dupas, Pascaline (2011). "Do Teenagers Respond to HIV Risk Information? Evidence From a Field Experiment in Kenya," *American Economic Journal: Applied Economics*, Vol. 3, No. (January), pages 1-34.

[51] Environmental Protection Agency (1990). "Regulatory Impact Analysis: Control of Sulfur and Aromatics Contents of On-Highway Diesel Fuel." Report EPA-420-R-90-103.

[52] Environmental Protection Agency (1993). "Final Regulatory Impact Analysis for Reformulated Gasoline." Report EPA-420-R-93-017.

[53] Environmental Protection Agency (2011). "The Benefits and Costs of the Clean Air Act from 1990 to 2020." Report.

[54] The Federal Trade Commission (2004). "The Petroleum Industry: Mergers, Structural Change and Antitrust Enforcement." Report. August 2004.

[55] The Federal Trade Commission (2006). "Investigation of Gasoline Price Manipulation and Post-Katrina Gasoline Price Increases." Report.

[56] The Federal Trade Commission (2006). "Gasoline Price Changes and the Petroleum Industry: An Update." Report. September 2011.

[57] The Federal Trade Commission (2010). *Horizontal Merger Guidelines*, Issued August 19, 2010.

[58] Figlio, David, and Maurice E. Lucas (2004). "What's in a Grade? School Report Cards and the Housing Market." *American Economic Review*, Vol. 94, No. 3, pages 591-604.

[59] Fischer, Carolyn (2005). "On the Importance of the Supply Side in Demand-Side Management." *Energy Economics*, Vol. 27, No. 1 (January), pages 165-180.

[60] Fischer, Carolyn (2010). "Imperfect Competition, Consumer Behavior, and the Provision of Fuel Efficiency in Light-Duty Vehicles." RFF Discussion Paper 10-60.

[61] Fowlie, Meredith, Mar Reguant, and Stephen Ryan (2014). "Market-based Environmental Regulation and the Evolution of Market Structure", forthcoming *Journal of Political Economy*

[62] General Accounting Office (2004). "Effects of Mergers and Market Concentration in the U.S. Petroleum Industry." Report to the Ranking Minority Member, Permanent Subcommittee on Investigations, Committee on Governmental Affairs, U.S. Senate. GAO-04-96.

[63] Gilbert, Richard and Justine Hastings (2005). "Market Power, Vertical Integration and the Wholesale Price of Gasoline." *The Journal of Industrial Economics*, Vol. LIII (4).

[64] Greenstone, Michael, Paul Oyer, and Annette Vissing-Jorgensen (2006). "Mandated Disclosure, Stock Returns, and the 1964 Securities Acts Amendments." *Quarterly Journal of Economics*, Vol. 121, No. 2, pages 399-460.

[65] Greenstone, Michael (2002). "The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufacturers." *Journal of Political Economy*, 110(6), pages 1175-1219.

[66] Grossman, Gene, and Carl Shapiro (1984). "Informative Advertising with Differentiated Products." *Review of Economic Studies*, Vol. 51, No. 1 (January), pages 63-81.

[67] Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica*, 50, 1029-1054.

[68] Hansen, Lars Peter and Kenneth J. Singleton (1982), "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." *Econometrica*, 50, 1029-1054.

[69] Hastings, Justine, and Jeffrey Weinstein (2008). "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." *Quarterly Journal of Economics*, Vol. 123, No. 4, pages 1373-1414.

[70] Hendricks, Ken, Preston McAfee, Michael Williams (2007), "Evaluating the Likely Competitive Effects of Horizontal and Vertical Mergers: A New Approach", *Antitrust Report*, (2007) Issue 2, pp. 33-40.

[71] Herberich, David, John List, and Michael Price (2011). "How Many Economists Does it Take to Change a Light Bulb? A Natural Field Experiment on Technology Adoption." Working Paper, University of Chicago (March).

[72] Hoffman, Florian, Roman Inderst, and Marco Ottaviani (2013). "Hypertargeting, Limited Attention, and Privacy: Implications for Marketing and Campaigning." Working Paper, Bocconi University (June).

[73] Houde, Sebastien (2014a). "How Consumers Respond to Environmental Certification and the Value of Energy Information." Working Paper, University of Maryland (March).

[74] Houde, Sebastien (2014b). "Bunching with the Stars: How Firms Respond to Environmental Certification." Working Paper, University of Maryland (February).

[75] Ichniowski, Casey, and Kathryn Shaw (2003). "Beyond Incentive Pay: Insiders' Estimates of the Value of Complementary Human Resource Management Practices." *Journal of Economic Perspectives*, Vol. 17, No. 1 (Winter), pages 155-180.

[76] Inderst, Roman, and Marco Ottaviani (2009). "Misselling Through Agents." *American Economic Review,* Vol. 99, No. 3, pages 883-908.

[77] Ishii, Joy (2004, "Compatibility, Competition, and Investment in Network Industries: ATM Networks in the Banking Industry." *Unpublished Ph.D. Thesis*, Harvard University.

[78] Iyengar, Sheena, and Mark Lepper (2000). "When Choice is Demotivating: Can One Desire Too Much of a Good Thing?" *Journal of Personality and Social Psychology*, Vol. 79, No. 6, pages 995-1006.

[79] Jessoe, Katrina and David Rapson (2014). "Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use." *American Economic Review*, Vol. 104, No. 4, pages 1417-1438.

[80] Jin, Ginger Zhe, and Phillip Leslie (2009). "Reputation Incentives for Restaurant Hygiene." *American Economic Journal: Microeconomics*, Vol. 1, No. 1 (February), pages 237-267.

[81] Jin, Ginger Zhe, and Alan Sorensen (2006). "Information and Consumer Choice: The Value of Publicized Health Plan Ratings." *Journal of Health Economics*, Vol. 25, No. 2, pages 248-275.

[82] Kallbekken, Steffen, Hakon Saelen, and Erlend Hermansen (2013). "Bridging the Energy Efficiency Gap: A Field Experiment on Lifetime Energy Costs and Household Appliances." *Journal of Consumer Policy*, Vol. 36, pages 1-16.

[83] Kahn, Matthew (2007). "Do Greens Drive Hummers? Environmental Ideology as a Determinant of Consumer Choice." *Journal of Environmental Economics and Management*, Vol. 54, No. 2 (September), pages 129-145.

[84] Kling, Jeffrey, Sendhil Mullainathan, Eldar Shafir, Lee Vermeulen, and Marian Wrobel (2012). "Comparison Friction: Experimental Evidence from Medicare Drug Plans." *Quarterly Journal of Economics.* Vol. 127, No. 1, pages 199-235.

[85] Larkin, Ian (2014) """The Cost of High-Powered Incentives: Employee Gaming in Enterprise Software Sales." *Journal of Labor Economics*, forthcoming.

[86] Lazear, Edward (2000). "Performance Pay and Productivity." *American Economic Review*, Vol. 90, No. 5 (December), pages 1346-1361.

[87] Leip, David (2013). "David Leip's Atlas of U.S. Presidential Elections." Available from http://uselectionatlas.org/

[88] Li, Shanjun, Joshua Linn, and Erich Muehlegger (2014), "Gasoline Taxes and Consumer Behavior.", American Economic Journal: Economic Policy, *forthcoming.*

[89] Lidderdale, Tancred, "Demand, Supply, and Price Outlook for Low-Sulfur Diesel Fuel." Energy Information Administration, Monthly Energy Review, August 1993.

[90] Mantrala, Murali, Sonke Albers, Fabio Caldieraro, Ove Jensen, Kissan Joseph, Manfred Krafft, Chakravarthi Narasimhan, Srinath Gopalakrishna, Andris Zoltners, Rajiv Lal, and Leonard Lodish (2010). "Sales Force Modeling: State of the Field and Research Agenda." *Marketing Letters*, Vol. 21, No. 3, pages 255-272.

[91] Marion, Justin and Erich Muehlegger (2008), "Measuring Illegal Activity and the Effects of Regulatory Innovation: Tax Evasion and the Dyeing of Untaxed Diesel." *Journal of Political Economy*, 114:6, pages 633-666.

[92] Milgrom, Paul (2008). "What the Seller Won't Tell You: Persuasion and Disclosure in Markets." *Journal of Economic Perspectives*, Vol. 22, No. 2 (April), pages 115-132.

[93] Millimet, Daniel L., Santanu Roy, and Aditi Sengupta (2009). "Environmental Regulations and Economic Activity: Influence on Market Structure." *Annual Review of Resource Economics*, Vol. 1, pages 99-117.

[94] Muehlegger, Erich (2004). "Product Differentiation in Gasoline Markets: A Discussion of Regional Content Regulations" *(working paper)*

[95] Muehlegger, Erich (2006). "Gasoline Price Spikes and Regional Gasoline Content Regulation: A Structural Approach." *(working paper)*

[96] Mullainathan, Sendhil, Markus Noeth, and Antoinette Schoar (2012). "The Market for Financial Advice: An Audit Study." NBER Working Paper No. 17929 (March).

[97] Nadel, Steven (2011). "Testimony of Steven Nadel, Executive Director, American Council for an Energy-Efficient Economy (ACEEE)." Before the Senate Energy and Natural Resources Committee, Hearing on Appliance Standards Regulation (March 10).

[98] Nagin, Daniel, James Rebitzer, Seth Sanders, and Lowell Taylor (2002). "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment." *American Economic Review*, Vol. 92, No. 4 (September), pages 850-873.

[99] NHTSA (National Highway Traffic Safety Administration) (2010). "Final Regulatory Impact Analysis: Corporate Average Fuel Economy for MY 2012-MY 2016 Passenger Cars and Light Trucks." Office of Regulatory Analysis and Evaluation, National Center for Statistics and Analysis (March).

[100] National Petroleum Council (1993). "U.S. Petroleum Refining: Meeting Requirements for Cleaner Fuels and Refineries." Report, August 1993.

[101] Newell, Richard, and Juha Siikamaki (2013). "Nudging Energy Efficiency Behavior: The Role of Information Labels." Resources for the Future Discussion Paper 13-17 (July).

[102] Pakes, Ariel, Jack Porter, Kate Ho, Joy Ishii (2015). "Moment Inequalities and Their Application." *Econometrica (forthcoming)*

[103] Pope, Devin (2009). "Reacting to Rankings: Evidence from 'America's Best Hospitals."' *Journal of Health Economics*, Vol. 28, No. 6, pages 1154-1165.

[104] Reiss, P. C., & Wolak, F. A. (2007). "Structural Econometric Modeling: Rationales and Examples from Industrial Organization." *Handbook of Econometrics*, 6, 4277-4415.

[105] Rosse, J.N. (1970). "Estimating Cost Function Parameters Without Using Cost Data: Illustrated Methodology". Econometrica 38 (2), 256–275

[106] Ryan, Stephen P. (2012). "The Costs of Environmental Regulation in a Concentrated Industry." *Econometrica*, Vol. 80, No. 3 (May), pages 1019-1062.

[107] Scanlon, Dennis, Michael Chernew, Catherine McLaughlin, and Gary Solon (2002). "The Impact of Health Plan Report Cards on Managed Care Enrollment." *Journal of Health Economics*, Vol. 21 No. 1, pages 19-41.

[108] Shearer, Bruce (2004). "Piece Rates, Fixed Wages, and Incentives: Evidence from a Field Experiment." *Review of Economic Studies*, Vol. 71, No. 2, pages 513-534.

[109] Sweeney, Richard L. and Thomas Wollmann (2015). "Two Period Strategies for Discrete Dynamic Games." *(working paper)*

[110] Ward, David, Christopher D Clark, Kimberly L Jensen, Steven T Yen, Clifford S Russell (2011). "Factors influencing willingness-to-pay for the ENERGY STAR label." *Energy Policy*, Vol. 39 No. 3, pages 1450-1458.

[111] Wollmann, Thomas (2014). "Trucks Without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles." *(working paper)*

[112] Zhang, Ker (2011). "The Effect of Sulfur Content Regulation on the U.S. Diesel Market." *(working paper)*

Appendix A

Appendix to Chapter 1

A.1 Summary of EIA data

Figure A.1: Description of EIA Data

Survey	Dates	Report Level	Description
Monthly Refinery Report (EIA-810)	Monthly 1986-2012	Refinery	Crude oil input volumes and characteristics. Balance of refined product supply, including beginning stocks, receipts, inputs, production, shipments, and ending stocks.
Annual Refinery Report (EIA-820)	Annual 1986-1995 1997 1999-2012	Refinery	Annual refinery energy purchased (fuel, electricity, and steam) for consumption at the refinery. Refinery receipts of crude oil by method of transportation. Current and projected capacities all atmospheric crude oil distillation and downstream units.
Refiners' Monthly Cost Report (EIA-14)	Monthly 2002-2012	Firm - PADD	Collects data on the weighted cost of crude oil booked into refineries in each PADD.
Refiners'/Gas Plant Operators' Monthly Petroleum Product Sales Report (EIA-782A)	Monthly 1986-2012	Firm - State	Price and volume data, by state, for 14 petroleum products at various retail and wholesale marketing categories. Reported by the universe of refiners and gas plant operators.
Monthly Report of Prime Supplier Sales of Petroleum Products Sold for Local Consumption (EIA-782C)	Monthly 1986-1990 1992-2012	Firm - State	Prime supplier sales volumes of selected petroleum products into the state of ultimate consumption. Reported by all refiners, gas plant operators, importers, petroleum product resellers, and petroleum product retailers that produce, import, or transport product across state boundaries.

Figure A.2: Wholesale Prices and Concentration by State

	Gasoline								Distillate							
	Pre-1993		1995-2003						Pre-1993		1995-2003					
	Conventional		Conventional		Reformulated		RFG	Price	High Sulfur		High Sulfur		Low Sulfur		LSD	Price
State	Price	HHI	Price	HHI	Price	HHI	Share	Diff.	Price	HHI	Price	HHI	Price	HHI	Share	Diff.
Padd 1 - East Coast																
CT	1.35	0.14			0.92	1.00			1.09	0.20	0.91	0.28	0.94	0.21	0.40	0.03
DC	1.42	0.31			0.00	0.00			1.07	0.46	0.88	0.81	1.03	0.38	0.48	0.11
DE	1.25	0.21			0.00	0.00			1.04	0.79	0.91	0.77	0.93	0.63	0.66	0.02
FL	1.24	0.07	1.04	0.11					1.06	0.10	0.90	0.14	0.93	0.11	0.70	0.03
GA	1.18	0.07	1.00	0.12					1.04	0.08	0.89	0.15	0.91	0.11	0.76	0.02
MA	1.34	0.15			0.80	0.93			1.09	0.22	0.90	0.28	0.95	0.23	0.37	0.05
MD	1.34	0.13	1.00	0.14	1.00	0.14	0.90	0.14	1.08	0.12	0.89	0.21	0.94	0.15	0.56	0.04
ME	1.23	0.20	1.02	0.25	1.02	0.25	0.24	-0.01	1.09	0.23	0.91	0.31	0.96	0.30	0.36	0.04
NC	1.17	0.08	0.99	0.12					1.04	0.09	0.89	0.11	0.91	0.13	0.66	0.02
NH	1.34	0.22	1.04	0.45	1.04	0.45	0.87	0.11	1.10	0.31	0.91	0.61	0.98	0.43	0.45	0.06
NJ	1.27	0.10	0.92	0.31	0.92	0.31	0.90	0.19	1.05	0.11	0.88	0.17	0.91	0.17	0.37	0.03
NY	1.35	0.14	1.03	0.20	1.03	0.20	0.51	0.18	1.09	0.12	0.92	0.16	0.96	0.16	0.34	0.04
PA	1.21	0.11	1.01	0.12	1.01	0.12	0.26	0.11	1.07	0.10	0.90	0.26	0.93	0.20	0.53	0.02
RI	1.31	0.16			0.89	1.00			1.08	0.34	0.91	0.32	0.95	0.38	0.35	0.04
SC	1.17	0.08	0.99	0.11					1.05	0.10	0.90	0.15	0.91	0.12	0.73	0.01
VA	1.23	0.08	0.98	0.12	0.98	0.12	0.57	0.14	1.05	0.08	0.89	0.12	0.92	0.13	0.63	0.03
VT	1.34	0.57	1.08	0.54					1.13	0.32	0.92	0.61	0.96	0.50	0.39	0.04
WV	1.23	0.23	1.02	0.19					1.09	0.23	0.95	0.52	0.95	0.26	0.47	0.00
Padd 2 - Midwest																
IA	1.18	0.08	1.03	0.11					1.08	0.12	0.95	0.21	0.97	0.14	0.86	0.02
IL	1.21	0.12	1.00	0.12	1.00	0.12	0.54	0.20	1.04	0.14	0.90	0.21	0.92	0.14	0.64	0.02
IN	1.17	0.12	1.02	0.18	1.02	0.18	0.14	0.12	1.06	0.13	0.93	0.19	0.93	0.15	0.71	0.00
KS	1.12	0.09	1.00	0.12					1.04	0.11	0.92	0.22	0.95	0.15	0.80	0.03
KY	1.21	0.16	1.01	0.20	1.01	0.20	0.26	0.10	1.08	0.17	0.93	0.30	0.94	0.31	0.57	0.02
MI	1.21	0.11	1.05	0.16					1.08	0.15	0.93	0.28	0.96	0.16	0.74	0.03
MN	1.20	0.11	1.08	0.13					1.09	0.12	0.96	0.23	0.98	0.19	0.70	0.02
MO	1.17	0.08	1.02	0.12	1.02	0.12	0.15	0.10	1.05	0.09	0.91	0.20	0.94	0.12	0.87	0.03
ND	1.21	0.15	1.05	0.21					1.11	0.21	1.00	0.69	1.01	0.31	0.67	0.01
NE	1.16	0.08	1.03	0.12					1.07	0.13	0.94	0.24	0.97	0.16	0.79	0.03
OH	1.20	0.17	1.05	0.19					1.10	0.17	0.95	0.31	0.96	0.24	0.66	0.00
OK	1.11	0.09	0.97	0.11					1.02	0.09	0.90	0.30	0.92	0.14	0.69	0.02
SD	1.19	0.08	1.05	0.11					1.10	0.11	0.96	0.27	1.00	0.14	0.91	0.03
TN	1.17	0.07	1.00	0.11					1.04	0.08	0.90	0.15	0.92	0.12	0.71	0.02
WI	1.18	0.09	1.02	0.13	1.02	0.13	0.29	0.10	1.07	0.13	0.93	0.18	0.95	0.21	0.73	0.02
Padd 3 - South																
AL	1.17	0.07	0.99	0.11					1.03	0.08	0.88	0.14	0.90	0.11	0.69	0.02
AR	1.15	0.07	0.99	0.10					1.04	0.09	0.88	0.23	0.92	0.12	0.65	0.04
LA	1.14	0.08	0.97	0.10	0.97	0.10	0.02	-0.02	1.00	0.12	0.84	0.18	0.88	0.12	0.31	0.04
MS	1.13	0.08	0.97	0.11					1.01	0.15	0.86	0.20	0.89	0.11	0.61	0.03
NM	1.23	0.09	1.07	0.12					1.12	0.11	0.96	0.43	1.00	0.17	0.92	0.04
TX	1.12	0.06	0.94	0.08	0.94	0.08	0.29	0.09	1.00	0.06	0.85	0.14	0.89	0.09	0.72	0.03
Padd 4 - Plains																
CO	1.19	0.10	1.07	0.13					1.08	0.12	0.96	0.48	0.99	0.15	0.84	0.03
ID	1.21	0.14	1.14	0.15					1.17	0.15	1.05	0.29	1.07	0.19	0.70	0.03
MT	1.20	0.17	1.13	0.23					1.12	0.20	1.00	0.76	1.05	0.24	0.98	0.07
UT	1.19	0.12	1.13	0.15					1.14	0.16	1.04	0.31	1.07	0.22	0.81	0.04
WY	1.21	0.12	1.10	0.14					1.12	0.20	1.02	0.84	1.03	0.15	0.79	0.00
Padd 5 - West Coast																
AZ	1.25	0.12	1.16	0.13	1.16	0.13	0.43	0.08	1.12	0.13	0.96	0.45	1.00	0.13	0.88	0.04
CA	1.27	0.10	1.07	0.31	1.07	0.31	0.93	0.16	1.04	0.11	0.98	0.46	1.03	0.17	0.96	0.05
NV	1.22	0.14	1.18	0.17					1.10	0.15	1.01	0.45	1.01	0.13	0.90	0.00
OR	1.23	0.14	1.14	0.18					1.06	0.21	0.97	0.44	0.99	0.20	0.64	0.02
WA	1.21	0.12	1.16	0.15					1.05	0.18	0.94	0.38	0.99	0.15	0.60	0.05

Notes: Prices are average monthly prices for resale ($2013 / gal). Markets with sales in less than 50% of the sample are excluded.

A.2 Crude price variable construction

Crude oil makes up over 80 percent of a refinery's variable costs. While crude oil quality and volumes are observed at the refinery level in the EIA data, crude oil prices are not. Prices are observed at the firm-PADD level beginning in 2004. In order to construct a proxy for refinery-level crude prices during the sample, I estimate a regression relating this firm-PADD level data to publicly available crude price series that are available for the entirety of the sample.[1] I then predict crude prices for each refinery in each month and use this variable when estimating the structural model.

Beginning in 2004, survey EIA-14 records the monthly total cost and volume of domestic (D) and imported (I) crude oil acquired across all refineries owned by a firm in each Petroleum Administration Defense District.[2] These costs are assumed to be a function of benchmark crude prices, PADD-level domestic crude prices and cost shifters, and a quality premium on API gravity. West Texas Intermediate and Brent spot prices serve as domestic and imported crude benchmark prices. Regional variation in domestic crude prices is captured by the EIA's cost of first purchase price (p^{cofp}) series, which reports the average price received by domestic oil producers in each PADD. The EIA reports the average landed cost of imported crude (p^{land}) by API gravity bin (b) beginning 1986.[3] Using the midpoint of each bin, I construct a price per API gravity degree premium variable for each month, $\zeta_b^I = \frac{p_b^{land} - p^{brent}}{API_b - API^{brent}}$. Domestic prices by API gravity bin are not available before 1994. Therefore the domestic API gravity premium is proxied with the average price premium of two domestic heavy crudes, Alaska North Slope an Gulf Heavy.

Given the available price data, firm-PADD crude price are modeled as follows,

$$\begin{aligned} p_{fr} &= \alpha_0 + s_{fr}(\alpha_{D0r} + \alpha_{D1r}p^{wti} + \alpha_{D2}\zeta^D\Delta_{fr}^D + \alpha_{D3}(p_r^{cofp} - p^{wti})) \\ &\quad + (1 - s_{fr})(\alpha_{I0r} + \alpha_{I1r}p^{brent} + \alpha_{I2}\zeta_{fr}^I\Delta_{fr}^I) + \alpha_2 s_{fr}\zeta^D\Delta_{fr}^D(1 - s_{fr})\zeta_{fr}^I\Delta_{fr}^I \end{aligned}$$

Where s_{rf} is the fraction of crude processed by firm f in PADD r that is domestic, and Δ is the difference in API gravity between this crude and the benchmark crude (i.e. $\Delta_{fr}^I = API_{fr} - API^{brent}$). The final term is an interaction between domestic and imported API gravity premiums to account for the fact than refineries only report average API gravity each month, rather than separate figures for domestic and imported crude streams.

[1] All of the crude price series discussed here can be downloaded at http://www.eia.gov/petroleum/data.cfm#prices

[2] The survey actually begins in 2002, but data was only collected at the national level until 2004.

[3] For example, the average cost of imported crude with API gravity between 20 and 25 degrees.

Table A.1: Crude Price Estimates

	Average Crude Price	
Constant	-0.159	(2.090)
Domestic - PADD 1	7.520***	(2.766)
Domestic - PADD 2	-2.461	(2.273)
Domestic - PADD 3	2.756	(2.284)
Domestic - PADD 4	4.453*	(2.429)
Domestic - PADD 5	1.207	(2.290)
COFP - PADD 1	1.021***	(0.207)
COFP - PADD 2	0.499***	(0.102)
COFP - PADD 3	1.139***	(0.0646)
COFP - PADD 4	0.962***	(0.115)
COFP - PADD 5	0.827***	(0.0417)
WTI - PADD 1	0.948***	(0.0207)
WTI - PADD 2	1.057***	(0.0104)
WTI - PADD 3	1.025***	(0.0101)
WTI - PADD 4	1.008***	(0.0148)
WTI - PADD 5	1.004***	(0.0103)
API Premium - Domestic	0.0554***	(0.0161)
Imported - PADD 1	3.071	(2.270)
Imported - PADD 2	12.11***	(2.343)
Imported - PADD 3	-2.336	(2.203)
Imported - PADD 4	12.87***	(2.355)
Imported - PADD 5	7.499***	(2.811)
Brent - PADD 1	0.920***	(0.0108)
Brent - PADD 2	0.717***	(0.0122)
Brent - PADD 3	0.979***	(0.00863)
Brent - PADD 4	0.687***	(0.0125)
Brent - PADD 5	0.924***	(0.0204)
Brent - PADD 6	0.999***	(0.0240)
API Premium - Imported	0.137***	(0.0106)
API Premium Interaction	-0.0319***	(0.00405)
N	8258	
r2	0.929	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

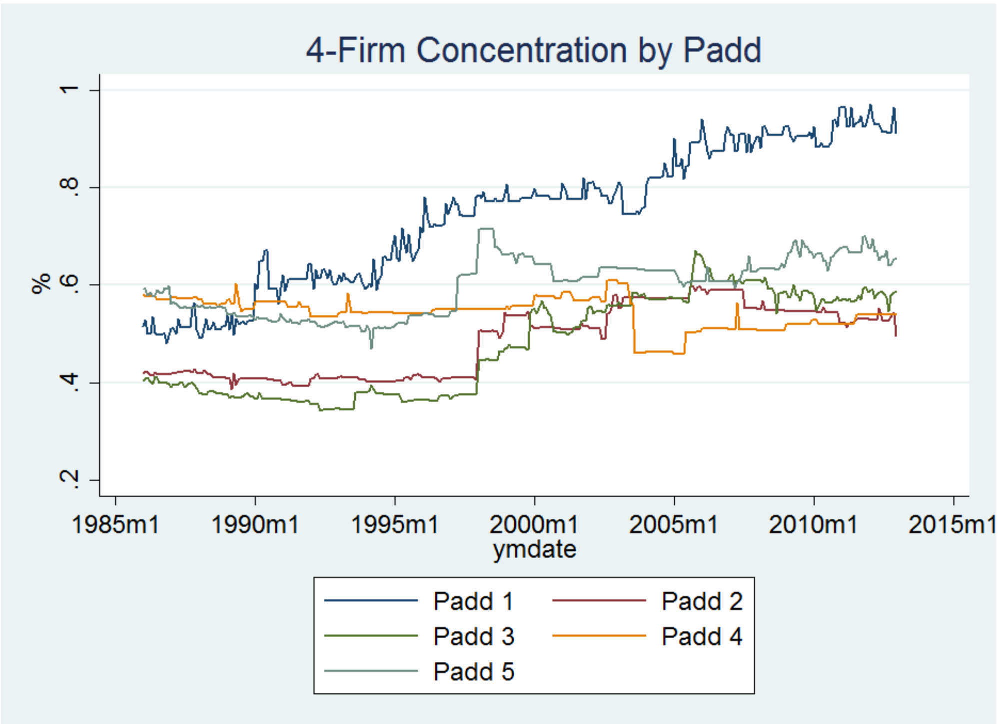
A.3 Market power and refinery utilization rates

In order to demonstrate the importance of market power in this setting, I run several versions of the following regression,

$$Utilization_{it} = \alpha_i + \gamma Capshare_{fi} + \epsilon_{it}$$

$Capshare_{ft}$ is the sum of distillation capacity across all refineries owned by firm f in the same PADD as refinery i divided by the total distillation capacity in that PADD during that month t. Regressions are run on the full set of data available, beginning in 1986 and ending in 2012. Ownership concentrations within a PADD vary considerably during this time period, as can be seen in Figure A.3. The primary driver of these changes was a wave of mergers and acquisitions seen in Table 1.3. In the IV regressions, I used the prior year's capacity at each refinery aggregated using the current year's ownership status. This instrument intuitively takes advantage of variation from ownership changes alone rather than increases in investment capacity within a given year.

Figure A.3: Changes in Capacity Concentration by PADD



Appendix B

Appendix to Chapter 3

B.1 Proofs of Theoretical Results

Proof of Proposition 3

The comparative static can be derived by simply writing out $q_{fE}(p_{fE}, \tilde{\theta}_f(p_{fE}, s_f)$ and taking derivatives. Consider firm 1:

$$q_{1E}(p_{1E}, \tilde{\theta}_1(p_{1E}, s_1) = \lambda\tilde{\theta}_1 \left[(1-\theta_2) + \theta_2 \frac{p_{2E} + t - p_{1E}}{2t}\right] \tag{B.1}$$

Substituting $\tilde{\theta}_1$ from Equation (3.2) gives:

$$q_{1E}(p_{1E}, \tilde{\theta}_1(p_{1E}, s_1) = \frac{s_1\lambda^2}{\alpha} \left[(1-\theta_2) + \theta_2 \frac{p_{2E} + t - p_{1E}}{2t}\right]^2 \tag{B.2}$$

Taking derivatives gives:

$$\begin{aligned} \frac{\partial q_{1E}}{\partial s_1} &= \frac{\lambda^2}{\alpha} \left[(1-\theta_2) + \theta_2 \frac{p_{2E}+t-p_{1E}}{2t}\right]^2 > 0 \\ \frac{\partial q_{1E}}{\partial p_{1E}} &= -\frac{s_1\lambda^2}{\alpha} \cdot \frac{\theta_2}{t} \left[(1-\theta_2) + \theta_2 \frac{p_{2E}+t-p_{1E}}{2t}\right] < 0 \\ \frac{\partial^2 q_{1E}}{\partial s_1 \partial p_{1E}} &= -\frac{\lambda^2}{\alpha} \cdot \frac{\theta_2}{t} \left[(1-\theta_2) + \theta_2 \frac{p_{2E}+t-p_{1E}}{2t}\right] < 0 \end{aligned} \tag{B.3}$$

Derivation of Market Equilibrium

Using Equation (3.4), we now solve for the firm's optimal prices and the optimal θ_f that the firm would like to induce.

We first solve for the symmetric equilibrium prices, which are identical to the standard textbook model. The first order condition (FOC) for p_{1I} is:

$$\frac{\partial \pi_1}{\partial p_{1I}} = d_I - \frac{1}{2t}(p_{1I} - c_I) = 0 \tag{B.4}$$

This is the usual FOC for the textbook two-firm Hotelling model: increasing p_{1I} increases revenues from the share d_I of consumers who purchase from firm 1, but quantity sold decreases by an amount inversely proportional to the transport cost.

The FOC for p_{1E} is:

$$\frac{\partial \pi_1}{\partial p_{1E}} = \lambda \left[\theta_1(1-\theta_2) + \theta_1\theta_2 \left(d_E - \frac{1}{2t}(p_{1E} - c_E) \right) \right] \tag{B.5}$$

This FOC is also standard to models like Grossman and Shapiro (1984). Raising p_{1E} increases revenues from the share of High type consumers that are informed only by firm 1, and none of these consumers substitute to firm 2 because they are unaware that firm 2 sells good E. Share $\theta_1\theta_2$ are informed by both firms, and increasing p_{1E} increases revenues from the share that buy from firm 1 but decreases that share by an amount inversely proportional to the transport cost.

Simplifying Equation (B.4) gives $p^*_{1I} = \frac{p^*_{2I}+c_I+t}{2}$. Imposing symmetry and solving gives $p^*_{1I} = p^*_{2I} = c_I + t$. Simplifying Equation (B.5) gives $p^*_{1E} = 2t\frac{(1-\theta_2)}{\theta_2} + 2td_E + c_E$. Imposing symmetry and solving gives $p^*_{1E} = p^*_{2E} = c_E + \left(\frac{2}{\theta} - 1\right)t$.

The FOC for θ_1 is:

$$\frac{\partial \pi_1}{\partial \theta_1} = \lambda \left\{ (p_{1E} - c_E) \left[(1-\theta_2) + d_E\theta_2 \right] - (1-\theta_2)d_I(p_{1I} - c_I) \right\} - \alpha\theta_1 \tag{B.6}$$

Increasing θ_1 induces more High type consumers to purchase good E but cannibalizes sales of good I. The first term inside the brackets represents increased profits from sales of good E, but the second term shows decreased profits from good I.

The FOCs for firm 2 are again symmetric after replacing d_I and d_E with $1-d_I$ and $1-d_E$, respectively. To solve for the symmetric equilibrium information provision level, set Equation (B.6) to zero and re-arrange:

$$(p_{1E} - c_E) \left[(1-\theta_2) + d_E\theta_2 \right] - (1-\theta_2)d_I(p_{1I} - c_I) = \frac{\alpha}{\lambda}\theta_1 \tag{B.7}$$

Imposing symmetry and simplifying gives

$$\left(1 - \frac{\theta^*}{2}\right)(p^*_E - c_E) - \frac{1}{2}(1-\theta^*)(p^*_I - c_I) = \frac{\alpha}{\lambda}\theta^* \tag{B.8}$$

Substituting in the equilibrium prices gives:

$$\left(1 - \frac{\theta^*}{2}\right) t \left(\frac{2}{\theta^*} - 1\right) - \frac{t}{2}(1-\theta^*) = \frac{\alpha}{\lambda}\theta^* \tag{B.9}$$

Upon simplification, we have:

$$\left(\frac{\alpha}{\lambda t} - 1\right)\theta^{*2} + \frac{5}{2}\theta^* - 2 = 0 \tag{B.10}$$

Of the two roots of this quadratic equation, only one gives systematically positive values of θ. This is $\theta^* = \frac{-5+\sqrt{32\frac{\alpha}{\lambda t}-7}}{4\left(\frac{\alpha}{\lambda t}-1\right)}$, which is decreasing in α and increasing in λ and t.

Socially-Optimal Information Provision

In the case of energy-using durables, the U.S. Department of Energy facilitates and encourages retailer information disclosure through the Energy Star Retail Partnership program. Imagine that the social planner could induce firms to undertake any symmetric amount of information provision. Taking the derivative of Equation (3.5), the socially optimal level of symmetric information provision would satisfy the following first order condition:

$$\frac{\partial SW}{\partial \theta_f} = \lambda(2 - 2\theta_f)(G_{HE} - G_I) - 2\alpha\theta_f \tag{B.11}$$

The first term is the gain from inducing an additional High type consumer to purchase good E. The second term is the marginal cost of information. The socially-optimal symmetric information provision level θ_f^+ is thus:

$$\theta_f^+ = \frac{G_{HE} - G_I}{G_{HE} - G_I + \frac{\alpha}{\lambda}}. \tag{B.12}$$

Optimal information provision is higher when the social gain from moving a High type consumer to good E is larger, when there are a larger share of High types in the population, and when advertising costs are smaller. Depending on the value of $G_{HE} - G_I$, θ_f^+ could be anywhere between 0 and 1, regardless of the values of other parameters. Thus, it is ambiguous whether the market equilibrium θ_f^* over- or under-provides information relative to θ_f^+. Notwithstanding, only θ_f^+ is a function of $G_{HE} - G_I$, while only θ_f^* is a function of t. Thus, the market is more likely to under-provide information when the social gain $G_{HE} - G_I$ from moving consumers to good E is large or when market power t is small.

Socially-Optimal Government-Provided Sales Incentive

Assume that $\theta_f^+ > \theta_f^*$ and assume that the government has provided a sales incentive s^+ to induce socially-optimal information provision θ^+. We know that the firm would not provide an additional positive sales incentive given that s^+ is already larger than the profit-maximizing level. The participation constraint again pins down w_1. Substituting the participation constraint into firm 1's profit function gives:

$$\pi_1(p_{1I}, p_{1E}) = q_{1I}(p_{1I}, \theta^+)\left[p_{1I} - c_I\right] + q_{1E}(p_{1E}, \theta^+)\left[p_{1E} - c_E - s^+\right] - \left[\overline{U} - s^+ q_{1E}(p_{1E}, \theta^+) + \frac{\alpha}{2}(\theta^+)^2\right]. \tag{B.13}$$

The price first-order conditions in Equations (B.4) and (B.5) above are unchanged, so the profit maximizing prices are determined by the same formulas. The one difference is that $\theta_f = \theta_f^+$, and since $\theta_f^+ > \theta_f^*$, p_{fE} decreases relative to the original market equilibrium.

B.2 Appendix: Additional Tables and Figures

Figure A.1: Experiment Website Go Screen

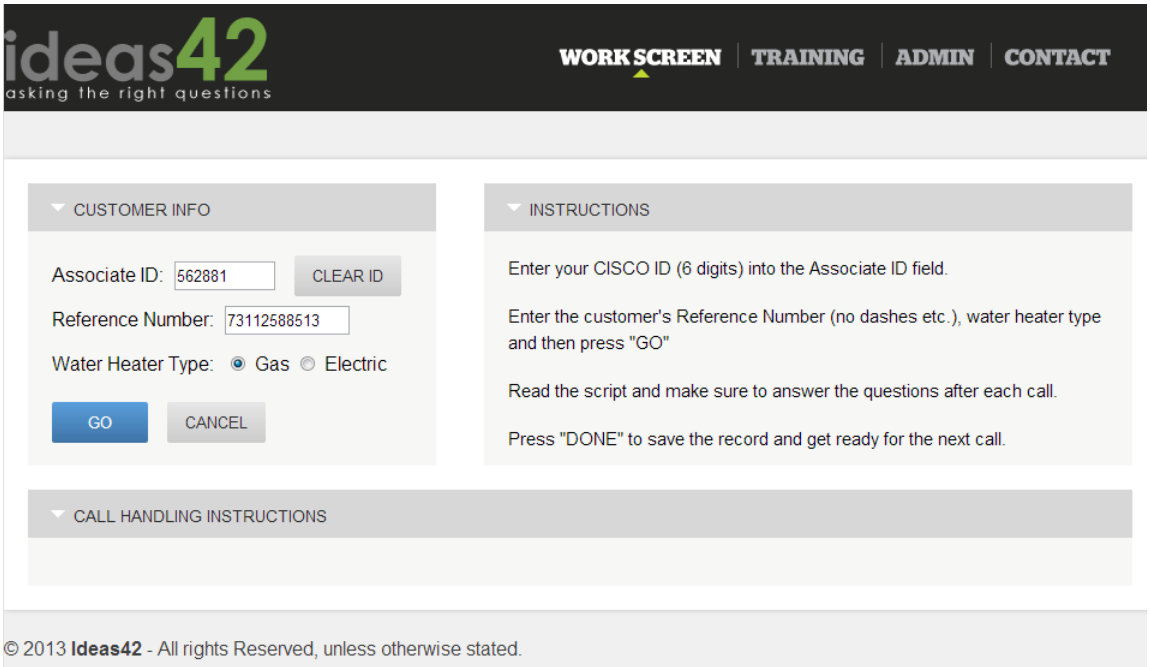


Figure A.2: Experiment Website $100 Rebate Screen

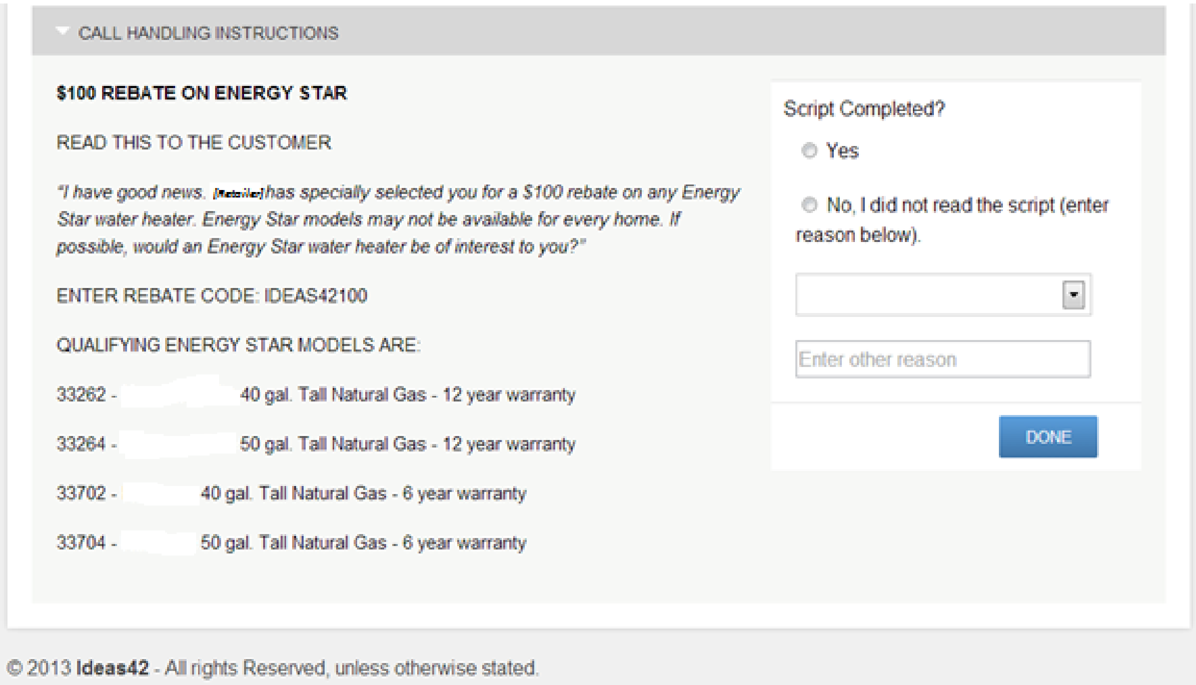


Table A.1: Test of Balance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Median Income	ln(Home Value)	College Grad	Age	Household Size	Democrat	Hybrid Share	Acxiom Green	Acxiom Enviro
1(Information Only)	1.192	0.039	0.009	-0.748	0.010	-0.013	0.066	-0.010	-0.013
	(1.185)	(0.037)	(0.018)	(0.568)	(0.067)	(0.015)	(0.051)	(0.020)	(0.015)
1(25 Rebate Only)	0.635	-0.004	0.013	-0.231	0.016	0.005	0.020	0.000	-0.009
	(0.814)	(0.026)	(0.013)	(0.398)	(0.044)	(0.010)	(0.035)	(0.014)	(0.011)
1(100 Rebate Only)	-1.499	-0.016	-0.029	-0.527	0.048	0.002	-0.052	0.009	0.009
	(1.110)	(0.037)	(0.019)	(0.564)	(0.066)	(0.015)	(0.045)	(0.021)	(0.016)
1(Info and 25 Rebate)	-1.162	-0.036	0.004	-0.299	-0.004	0.011	0.010	-0.025	-0.024
	(1.161)	(0.038)	(0.018)	(0.592)	(0.063)	(0.015)	(0.049)	(0.020)	(0.015)
1(Info and 100 Rebate)	0.520	0.057	0.013	-1.190	0.128	0.054	-0.038	0.021	0.001
	(2.158)	(0.068)	(0.036)	(1.116)	(0.117)	(0.027)**	(0.084)	(0.039)	(0.029)
1(Spiff Only)	0.053	-0.010	-0.008	-0.087	-0.058	0.007	-0.046	-0.016	-0.010
	(1.178)	(0.037)	(0.018)	(0.572)	(0.063)	(0.015)	(0.047)	(0.020)	(0.015)
1(Spiff and 25 Rebate)	-3.218	-0.024	0.015	0.038	0.175	-0.008	0.052	0.088	0.017
	(2.820)	(0.097)	(0.047)	(1.505)	(0.157)	(0.041)	(0.139)	(0.057)	(0.043)
1(Spiff and 100 Rebate)	-0.615	0.047	-0.032	-1.707	-0.105	-0.091	0.096	0.024	-0.002
	(4.061)	(0.125)	(0.078)	(2.006)	(0.269)	(0.067)	(0.219)	(0.082)	(0.061)
N	8,255	7,810	8,268	8,269	8,268	7,856	7,851	7,832	7,832
F-test p-Value	.426	.817	.642	.89	.777	.417	.583	.656	.820

Notes: This table reports regression of zip code level purchaser demographics on treatment assignment. All regressions contain agent and period indicator variables. Robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A.2: Treatment Effects: Alternative Estimates

	(1)	(2)	(3)
Outcome:	1(Sale)	1(Sale)	1(EStar)
1(100 Rebate)	-0.020	0.013	0.064
	(0.013)	(0.020)	(0.024)***
1(25 Rebate)	-0.027	-0.001	0.003
	(0.009)***	(0.013)	(0.010)
1(Information)	0.002	-0.021	0.006
	(0.011)	(0.016)	(0.013)
1(Spiff)	0.006	0.011	-0.000
	(0.015)	(0.019)	(0.013)
1(Spiff and 25 Rebate)	-0.027	0.044	-0.015
	(0.041)	(0.053)	(0.010)
1(Spiff and 100 Rebate)	-0.008	0.136	0.405
	(0.064)	(0.092)	(0.221)*
R^2	0.02	0.02	0.01
N	16,117	7,230	23,347
Dep. Var. Control Mean	.366	.358	.009
Regression Type:	ITT	ITT	Scaled ITT
Sample:	Non-Compliant Agents	Compliant Agents	All Consumers

Notes: This table reports the estimates of Equation (3.11), as alternative specifications to Table 3.6. Columns 1 and 2 are comparable to column 1 of Table 3.6, except splitting the sample to non-compliant agents (the lower two compliance groups) and compliant agents (the higher two compliance groups). Column 3 is comparable to column 6 of Table 3.6, except that G_{ia} is based on each compliance group's average difference in *Mentioned E-Star* between treatment and control calls. All regressions contain agent and period indicator variables. Robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A.3: Treatment Effects in the Sample Purchasing Substitutable Models

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	1(EStar)	1(EStar)	1(EStar)	1(EStar)	1(EStar)	1(EStar)
1(100 Rebate)	0.023	0.042	0.100	0.021	0.042	0.098
	(0.010)**	(0.017)**	(0.035)***	(0.011)**	(0.019)**	(0.038)***
1(25 Rebate)	0.006	0.013	0.010	0.007	0.014	0.009
	(0.005)	(0.009)	(0.015)	(0.006)	(0.010)	(0.016)
1(Information)	0.001	-0.000	0.016	-0.001	-0.003	0.014
	(0.006)	(0.012)	(0.020)	(0.007)	(0.013)	(0.022)
1(Spiff)	-0.010		-0.001	-0.009		-0.003
	(0.008)		(0.019)	(0.008)		(0.021)
1(Spiff and 25 Rebate)	-0.020		-0.023	-0.025		-0.026
	(0.009)**		(0.018)	(0.013)**		(0.028)
1(Spiff and 100 Rebate)	0.198		0.536	0.254		0.614
	(0.093)**		(0.242)**	(0.108)**		(0.260)**
R^2	0.04	0.04	0.04	0.12	0.13	0.13
N	6,000	5,180	6,000	6,000	5,180	6,000
Dep. Var. Control Mean	.033	.033	.033	.033	.033	.033
Regression Type:	ITT	Self-Report IV	Scaled ITT	ITT	Self-Report IV	Scaled ITT
Controls:	Base	Base	Base	Zip3	Zip3	Zip3

Notes: This table reports the estimates of Equation (3.11) in the sample of consumers that purchased substitutable models. In the Scaled ITT, the average compliance G_{ia} is calculated only within the sample of calls by consumers who purchase substitutable models. All regressions contain agent and period indicator variables, and columns 4-6 additionally include a vector of three-digit zip code indicators. Robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively.

Table A.4: Treatment Effects: Probit Estimates

Column in Table 6:	(1)	(3)	(4)	(6)
Dependent Variable:	1(Sale)	1(Sale)	1(EStar)	1(EStar)
1(100 Rebate)	-0.028	0.079	0.185	1.098
	(0.030)	(0.121)	(0.079)**	(0.295)***
1(25 Rebate)	-0.052	-0.044	0.022	0.126
	(0.019)***	(0.078)	(0.060)	(0.254)
1(Information)	-0.014	-0.160	0.029	0.181
	(0.024)	(0.101)	(0.074)	(0.323)
1(Spiff)	0.020	0.065	-0.137	-0.034
	(0.032)	(0.112)	(0.110)	(0.376)
1(Spiff and 25 Rebate)	0.001	0.142		
	(0.087)	(0.288)		
1(Spiff and 100 Rebate)	0.108	0.519	0.939	3.055
	(0.138)	(0.490)	(0.263)***	(0.711)***
N	23,338	23,338	21,613	21,699
Dep. Var. Control Mean	.364	.364	.009	.009
Regression Type:	ITT	Scaled ITT	ITT	Scaled ITT

Notes: This table reports probit estimates of Equation (3.11), as a robustness check to Table 3.6. The top row lists the column number corresponding to Table 3.6; columns 2 and 5 cannot be estimated in probit because they use instrumental variables with a binary dependent variable and binary instrument. The sample sizes differ slightly because the probit estimator drops observations with indicators that perfectly predict failure. This drops several μ_a coefficients in all columns and the 1(Spiff and 25 Rebate) indicator in the right two columns. All regressions contain agent and period indicator variables. Robust standard errors in parenthesis. *, **, ***: Statistically significant with 90, 95, and 99 percent confidence, respectively.